

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 23:34:02 ; Search time 88.5 Seconds
(without alignments)
2459.137 Million cell updates/sec

Title: US-10-728-947-2

Perfect score: 2155

Sequence: 1 VAEIMHVFAREILDSRGNPT.....QLLGDAGVVAGSAFPRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2152	99.9	425	1 ENO_CORGL	Q8nr61 corynebacte
2	2029	94.2	425	1 ENO_CORF	Q8fq87 corynebacte
3	1917	89.0	425	2 Q6NT61	Q6ni61 corynebacte
4	1599.5	74.2	426	1 ENO1_STRCO	Q8f2d3 streptomyce
5	1571.5	72.9	426	2 Q6ADR6	Q6adr6 leifsonia x
6	1558.5	72.3	429	2 Q741U7	Q741u7 mycobacteri
7	1554.5	72.1	428	2 Q82HH5	Q82hh5 streptomyce
8	1551.5	72.0	429	1 ENO_MYCTU	P96377 mycobacteri
9	1551.5	72.0	429	2 Q7U0U6	Q7u0u6 mycobacteri
10	1508.5	70.0	426	2 Q6AAB8	Q6aab8 propionibac
11	1500.5	69.6	429	1 ENO_MYCLE	Q9cd42 mycobacteri
12	1436	66.6	428	2 Q83H73	Q83h73 tropheryma
13	1436	66.6	446	2 Q83FF7	Q83ff7 tropheryma
14	1418	65.8	424	1 ENO_AGRT5	Q8ufh1 agrobacteri
15	1412	65.5	429	1 ENO_THETN	Q8r967 thermoanaer
16	1408.5	65.4	426	1 ENO_CAUCR	Q8a7j9 caulobacter
17	1406	65.2	429	1 ENO_BACHD	Q9k7i7 bacillus ha
18	1404.5	65.2	429	2 Q84FY9	Q84fy9 methyllobact
19	1402	65.1	425	1 ENO_BRUSU	Q8q0g3 brucella su
20	1400	65.0	424	1 ENO_RHILO	Q98m23 rhizobium l
21	1395	64.7	425	1 ENO_BRUME	Q8vfh0 brucella m
22	1394.5	64.7	430	2 Q65EN2	Q65en2 bacillus li
23	1393	64.6	424	1 ENO_RHIME	Q92q98 rhizobium m
24	1381.5	64.1	430	1 ENO_BACSU	P37869 bacillus su
25	1380.5	64.1	431	1 ENO_CLOPE	Q8xku4 clostridium
26	1379	64.0	431	2 Q815K8	Q815k8 bacillus ce
27	1377	63.9	431	2 Q81X78	Q81x78 bacillus an
28	1377	63.9	431	2 Q6HBF3	Q6hbf3 bacillus th
29	1375	63.8	431	2 Q63IM2	Q63lm2 bacillus ce
30	1374	63.8	428	2 Q74AR6	Q74ar6 geobacter s
31	1373	63.7	432	2 Q7VDY0	Q7vdy0 prochloroco

32 1370.5 63.6 427 1 ENO_RALSO
33 1366.5 63.4 430 2 Q7U3T1
34 1366 63.4 431 2 Q72XV5
35 1365.5 63.4 430 2 Q67SV9
36 1360.5 63.1 430 1 ENO_LISIN
37 1360.5 63.1 430 2 Q71WX1
38 1360.5 63.1 430 2 Q898R0
39 1360.5 63.1 431 2 Q898R0
40 1359.5 63.1 429 1 ENO_OCEIH
41 1359.5 63.1 434 2 Q8GDZ5
42 1358.5 63.0 431 1 ENO_CLOAB
43 1353 62.8 430 2 Q7V377
44 1351 62.7 431 2 Q7V483
45 1346 62.5 432 1 ENO_ENTFA

ALIGNMENTS

RESULT 1
ID ENO_CORGL STANDARD; PRT; 425 AA.
AC Q8NR61; 2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
glycerate hydro-lyase).
DE Glycerate hydro-lyase).
GN Name=eno; OrderedLocusNames=Cgl0974, cgl111;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
Burkovski A., Busch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
and its impact on the production of L-aspartate-derived amino acids
and vitamins.";
RL J. Biotechnol. 104:5-25 (2003).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
H(2O).
CC -I- COFACTOR: Magnesium is required for catalysis and for stabilizing
the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib.ch/>)
or send an email to license@ebi-sib.ch.
CC -----
DR EMBL; AP005277; BAB98367.1; -.
DR EMBL; BX927150; CAF19681.1; -.
DR HSSP; P08324; 1E9I.

Q8Y0B5 ralstonia s
Q7u3t1 synechococc
Q72xv5 bacillus ce
Q67sv9 symbiobacte
P64075 listeria in
P64074 listeria mo
Q71wx1 listeria mo
Q898r0 clostridium
Q8enp5 oceanobacil
Q8gdz5 heliobacill
Q97152 clostridium
Q7v483 prochloroco
Q9k596 enterococcu


```

RESULT 3
O6NI61 PRELIMINARY; PRT; 425 AA.
AC Q6NI61;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name=eno; OrderedLocusNames=DIP0917;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdeno-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,
RA De Zuyasa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523 (2003).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
H(2)O.
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
DR EMBL: BX248356; CAE49433.1; -.
DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:0016829; F:lyase activity; IEA.
DR GO: GO:0000287; F:magnesium ion binding; IEA.
DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; F:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; Enolase_C; 1.
DR Pfam: PF03952; Enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
SK Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 425 AA; 45029 MW; 85BF477EE7ECDCA6 CRC64;

Query Match 89.0%; Score 1917; DB 2; Length 425;
Best Local Similarity 88.0%; Pred. No. 3.1e-105;
Matches 374; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

QY 1 VAETMHVFARILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGERYLG 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MADIMHVFARILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGERYLG 60

QY 61 KGVLLKAVENNEETIGDELAGLEADDQRLIDAMIKLDGTANKSRILGANAILGVSMAYAKA 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 KGVLLNAVNNVEETADATAGAEADDQRLIDQAMIALDGTENKSRILGANAILGVSIAYAKA 120

QY 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGAHADSGVDVQEPMIAPIGAEFTSEALRN 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AAESAGLPLFRYIGGPNHVLVPVPMNIIINGAHADSGVDVQEPMIAPIGAEFTSEALRM 180

QY 181 GAETHYALKSVIKKGLSTGLDGGGFPSPVSGTREALDLIVEAIEKAGFTPGKDIALAL 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GAETHYLSKSVIKKGLSTGLDGGGFPSPVSGTREALDLIVEAIEKAGFKPGADIALAL 240

QY 241 DVASSEFFKDCYTHFEGGQHSAAEMANVYAEVLDAVPVTSIEDPLQEDDWEGYTALTAI 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 DVASSEFFKDCYTHFEGGQHSAAEMANVYAEVLDAVPVTSIEDPLQEDDWEGYTALTAI 300

```



```

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
DR EMBL; AE017230; AA03307.1; -.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR PRINTS; PR00148; ENOLASE.
DR PRODOM; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 429 AA; 44873 MW; C5B7C046F01D703 CRC64;

Query Match 72.3%; Score 1558.5; DB 2; Length 429;
Best Local Similarity 72.9%; Pred. No. 4.5e-84;
Matches 307; Conservative 46; Mismatches 67; Indels 1; Gaps 1;

Qy 4 IMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGVRLGKGV 63
Db 4 IEQVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGVRLGKGV 63

Qy 64 LKAVENNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGTFARAAVPSGASTGVHEAHELRDGGVRLGKGV 63

Qy 124 SAGLPLFRYIGGPNNAHLVPPVPMNIINGGAHADSGVDVQEPFMIAPIGAEFTSEALRNGAE 183
Db 124 SAEPLFRYIGGPNNAHLVPPVPMNILINGGAHADTAVDIEQFVAPIGAPSPAEALRWGAE 183

Qy 184 VYHAKSVIKEKGLSTGLDGGGAPSVGSTREALDLIVEAIEKAGFTPGKDIALADVA 243
Db 184 VYHLSKSVLKEGSLSTGLDGGGAPDVAGTAAALDLIGRAIESAGFKLGIDVALADAA 243

Qy 244 SSEFFKDKGT-YHFEFGQHSAAEMANVYAEVLDVAYPIVSIEDPQEDDWEYTNLTATIGD 302
Db 244 ATEFYSDGTGYKFGSTRTAQMAEFYAGLLGAYPLVSIEDPLSDDDWDGWAALTASIGD 303

Qy 303 KVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTLTTFDAVDMAHRAGYTSMWSH 362
Db 304 RVQLVGDDVFTNPERLEEGYEKGVANALLVKVNIQIGLTLTDLDAVALAHSHGYRTWWSH 363

Qy 363 RSGTETDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 422
Db 364 RSGTETDTTIADLAVAGSGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRLAPPR 423

Qy 423 F 423
Db 424 F 424

RESULT 7
Q82HH5 PRELIMINARY; PRT; 428 AA.
AC Q82HH5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative enolase.
GN Name=eno; OrderedLocusNames=SAV3533;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.2114333198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
DR EMBL; AP005035; BAC71245.1; -.
DR HSP; Q9NDH8; 10EP.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR PRINTS; PR00148; ENOLASE.
DR PRODOM; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 428 AA; 45869 MW; 095A3976F6C5F70D CRC64;

Query Match 72.1%; Score 1554.5; DB 2; Length 428;
Best Local Similarity 73.7%; Pred. No. 7.7e-84;
Matches 309; Conservative 42; Mismatches 67; Indels 1; Gaps 1;

Qy 7 VFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGKGVLK 65
Db 7 VVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDPNRYQKGVK 66

Qy 66 AVENVNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAADS 125
Db 67 AVLAVIEIQIGPELVGYDATEQRLIDQAMFOLDLTDNKGSLGANAILGVSLAVAAAS 126

Qy 126 GLPLFRYIGGPNNAHLVPPVPMNIINGGAHADSGVDVQEPFMIAPIGAEFTSEALRNGAEVY 185
Db 127 DLPLFRYIGGPNNAHLVPPVPMNILINGGAHADSDNYDIEQFMIAPIGAEFTSEALRNGAEVY 186

Qy 186 HALKSVIKEKGLSTGLDGGGAPSVGSTREALDLIVEAIEKAGFTPGKDIALADVA 245
Db 187 HTLKKVLTGSLTGLDGGGAPNLSNRALDLIIETAKQGYIPGEQIALADVA 246

Qy 246 EFFKDGTYHFEFGQHSAAEMANVYAEVLDVAYPIVSIEDPQEDDWEYTNLTATIGDKVQ 305
Db 247 EFKDGTYHFEFGKSRSAEWEYEEVLSVPLVSIEDPQEDDWEYTNLTATIGDKVQ 306

Qy 306 IVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTLTTFDAVDMAHRAGYTSMWSHRS 365
Db 307 IVGDDFFVTNPERLARGIEGSANALLVKVNIQIGLTLTTFDAVDMAHRAGYTSMWSHRS 366

```


CC -|- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H₂O.
 CC -|- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).
 CC -|- PATHWAY: Glycolysis.
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the enolase family.
 DR EMBL: BX248337; CAD93911.1; -.
 DR HSP: P08324; I891.
 DR GO: GO:000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 DR GO: GO:0000287; F:magnesium ion binding; IEA.
 DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR000941; Enolase.
 DR Pfam: PF00113; Enolase_C; 1.
 DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 429 AA; 44862 MW; F808FE80D891C9A CRC64;

Query Match 72.0%; Score 1551.5; DB 2; Length 429;
 Best Local Similarity 73.2%; Pred. No. 1.2e-83;
 Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNPTVEAEVFLDDSGHGVAGVPSCASTGVHAEHLRGGDRYLKGV 63
 DB 4 IEQVGAREILDSRGNPTVEAEVFLDDSGHGVAGVPSCASTGVHAEHLRGGDRYLKGV 63
 QY 64 LKAVENNEEIGDELGLADDDQRLIDEAMIKLDGTANKSRGLGANATILGVSMVAKAAD 123
 DB 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGTPKRLGGNATILGVSLAVAKAAD 123
 QY 124 SAGLPLFRYIGGNNAHLVPPMNIINGGAHDSGVVDQEFMIAPIGAEFTFSEALRNGAE 183
 DB 124 SAEPLFRYVGNNAHLVPPMNIINGGAHADTAVDIQEFVAPIGAPSFVEALRWGAE 183
 QY 184 VYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALADVA 243
 DB 184 VYHALKSVLKEGLSTGLDEGGFAPSVAGTAAALDLISRAIESAGLRPGADVALADAA 243
 QY 244 SREFFKDGTYHFEFGQHSAAEMANVYAEIVDAPYIVSIEDPDEDDWEGYTNLTATIGD 302
 DB 244 ATEFFDTGTVGTEGTRTDQMTFETAGLGLAYPLVSIEDPDEDDWEGYTNLTATIGD 303
 QY 303 KQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGRYTSMMSH 362
 DB 304 RVQIVGDDIFVTNPERLEEGIERGVANALLVKVNIQIGTLTETLDVTLAHGGYRTMISH 363
 QY 363 RSGETEDTTIADLAVLNCCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAPFR 422
 DB 364 RSGETEDTTIADLAVLNCCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAPFR 423
 QY 423 F 423
 DB 424 F 424

RESULT 10
 Q6AAB8 PRELIMINARY; PRT; 426 AA.
 ID Q6AAB8
 AC Q6AAB8; (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Enolase (EC 4.2.1.11)
 GN OrderedLocustNames=PPA0545;
 OS Propionibacterium acnes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacteriaceae; Propionibacterium.

OX NCBI_TaxID=1747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KPA171202 / DSM 16379;
 RX PubMed=15286373; DOI=10.1126/science.1100330;
 RA Brueggemann H., Henne A., Hofer F., Liesegang H., Wierer A.,
 RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.,
 RT "The complete genome sequence of Propionibacterium acnes, a commensal of human skin".
 RL Science 305:671-673(2004).
 CC -|- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H₂O.
 CC -|- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).
 CC -|- PATHWAY: Glycolysis.
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the enolase family.
 DR EMBL: AE017283; AAT82298.1; -.
 DR GO: GO:000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR000941; Enolase.
 DR Pfam: PF00113; Enolase_C; 1.
 DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 426 AA; 45530 MW; 39FDOE9A2941123F CRC64;

Query Match 70.0%; Score 1508.5; DB 2; Length 426;
 Best Local Similarity 68.9%; Pred. No. 4e-81;
 Matches 293; Conservative 59; Mismatches 72; Indels 1; Gaps 1;

QY 1 VAMHVFAREILDSRGNPTVEAEVFLDDSGHGVAGVPSCASTGVHAEHLRGGDRYL 59
 DB 1 MATIEAREILDSRGNPTVEAEVFLDDSGHGVAGVPSCASTGVHAEHLRGGDRYL 60
 QY 60 KGKVLKAVENNEEIGDELGLADDDQRLIDEAMIKLDGTANKSRGLGANATILGVSMVAK 119
 DB 61 KGKVLKAVENNEEIGDELGLADDDQRLIDEAMIKLDGTANKSRGLGANATILGVSMVAK 120
 QY 120 AAADSGLPLFRYIGGNNAHLVPPMNIINGGAHDSGVVDQEFMIAPIGAEFTFSEALR 179
 DB 121 AAADSGLPLFRYIGGNNAHLVPPMNIINGGAHDSGVVDQEFMIAPIGAEFTFSEALR 180
 QY 180 NGAEVYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALA 239
 DB 181 WGAHVYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALA 240
 QY 240 LDVASSEFFKDGTYHFEFGQHSAAEMANVYAEIVDAPYIVSIEDPDEDDWEGYTNLTAT 299
 DB 241 LDVASSEFFKDGTYHFEFGQHSAAEMANVYAEIVDAPYIVSIEDPDEDDWEGYTNLTAT 300
 QY 300 IGDVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGRYTSM 359
 DB 301 LGSEKIQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGRYTSM 360
 QY 360 MSHRSGETEDTTIADLAVLNCCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRS 419
 DB 361 MSHRSGETEDTTIADLAVLNCCQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRS 420
 QY 420 FPRFQ 424
 DB 421 FPRFQ 425

RESULT 11
 ENO_MYCLE STANDARD; PRT; 429 AA.
 ID ENO_MYCLE


```

Query Match      66.6%; Score 1436; DB 2; Length 428;
Best Local Similarity 67.9%; Pred. No. 7.7e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

QY 7 VFAREILDSRGNPTEAEVFLDDSGHVAGVPSGASTGVHEAHLRD-GGDRYLKGVLK 65
DB 7 VHARQILDSRGTPTEVEVTELEDGSMGRSAVPSGASTGAFAEHLRDQDNNEYLKGVTR 66
QY 66 AVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRIGCANAILGVSMVAKAAADSA 125
DB 67 AVRSVNSEIAPVLIGFPAFQVGLDHRMIELDGTNNKSRIGCANAILGVSLASASAAARA 126
QY 126 GLPLFRYIGGNPAHVLFPVPMNIIINGGAHDSGVDOEFMTIAPICAEFTFSEALRNGAEVY 185
DB 127 DLSLFRYLGGPSSRILFPVPMNIIINGGAHADTGVDIQEFMLPVGARSFSESRLWGCEVY 186
QY 186 HALKSIVKEKLSGLGDEGGFAPSVGSTREALDLIVEAIKAGFTPGKDTALALDVASS 245
DB 187 HSLKVQLRESGLSGLDGEGFAPALRSNRRTALDLILSAIEKAGFSPGIDIVLALDIAAS 246
QY 246 EEFK-DGTYHFEFGQHSAAEMANVYAEI VDAYPVSTIEDPQEDDWEQYTNLTATIGDKV 304
DB 247 EFCAPGHYRPEGKDI TSDELISYEGLLSYPLVSIEDPQDDWEGYRTLTTHIGDRV 306
QY 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVQNIQIGTLTETFDVDMHRAAGYTTSMMSHRS 364
DB 307 QIVGDDLVFVNTSRLSRGIQSGVANSILVKVQNIQIGTLTETFDVDMAAKGGYTAVLSHRS 366
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPFRPQ 424
DB 367 GETEDTTIADMAVATNCCQIKTGAPARGERTAKYNQLLRIEKLGRSARYAGWLSYPRWQ 426
QY 425 G 425
DB 427 G 427

RESULT 13
Q83FF7 PRELIMINARY; PRT; 446 AA.
AC Q83FF7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name=eno; OrderedLocusNames=TW783;
OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=203267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Twist;
RA MEDLINE=22784088; PubMed=12902375;
RA Raoult D., Ogata H., Audic S., Robert C., Suhre K., Drancourt M.,
RA Claverie J.-M.;
RT "Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a
RT reduced genome.";
RL Genome Res. 13:1800-1809 (2003).
CC -|- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -|- COPACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -|- PATHWAY: Glycolysis.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the enolase family.
DR EMBL: AE016852; AA044880.1; -.
DR HSP: Q9NDH8; 10EP.
DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:0016829; F:lyase activity; IEA.
DR GO: GO:0000287; F:magnesium ion binding; IEA.
DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; Enolase_C; 1.
DR Pfam: PF03952; Enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 446 AA; 48141 MW; 786E73F99F94A6DA CRC64;

Query Match      66.6%; Score 1436; DB 2; Length 446;
Best Local Similarity 67.9%; Pred. No. 8e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

QY 7 VFAREILDSRGNPTEAEVFLDDSGHVAGVPSGASTGVHEAHLRD-GGDRYLKGVLK 65
DB 25 VHARQILDSRGTPTEVEVTELEDGSMGRSAVPSGASTGAFAEHLRDQDNNEYLKGVTR 84
QY 66 AVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRIGCANAILGVSMVAKAAADSA 125
DB 85 AVRSVNSEIAPVLIGFPAFQVGLDHRMIELDGTNNKSRIGCANAILGVSLASASAAARA 144
QY 126 GLPLFRYIGGNPAHVLFPVPMNIIINGGAHDSGVDOEFMTIAPICAEFTFSEALRNGAEVY 185
DB 145 DLSLFRYLGGPSSRILFPVPMNIIINGGAHADTGVDIQEFMLPVGARSFSESRLWGCEVY 204
QY 186 HALKSIVKEKLSGLGDEGGFAPSVGSTREALDLIVEAIKAGFTPGKDTALALDVASS 245
DB 205 HSLKVQLRESGLSGLDGEGFAPALRSNRRTALDLILSAIEKAGFSPGIDIVLALDIAAS 264
QY 246 EEFK-DGTYHFEFGQHSAAEMANVYAEI VDAYPVSTIEDPQEDDWEQYTNLTATIGDKV 304
DB 265 EFCAPGHYRPEGKDI TSDELISYEGLLSYPLVSIEDPQDDWEGYRTLTTHIGDRV 324
QY 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVQNIQIGTLTETFDVDMHRAAGYTTSMMSHRS 364
DB 325 QIVGDDLVFVNTSRLSRGIQSGVANSILVKVQNIQIGTLTETFDVDMAAKGGYTAVLSHRS 384
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPFRPQ 424
DB 385 GETEDTTIADMAVATNCCQIKTGAPARGERTAKYNQLLRIEKLGRSARYAGWLSYPRWQ 444
QY 425 G 425
DB 445 G 445

RESULT 14
ENO_AGRT5 STANDARD; PRT; 424 AA.
AC Q8UFH1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocusNames=Atu1426, AGR_C.2631;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,

```

RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.,
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58".
RL Science 294:2317-2323 (2001).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houtmei K., Gordon J., Vaudin M., Hartchouk O., Epp A., Liu F.,
RA Wollan C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.,
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58".
RL Science 294:2323-2328 (2001).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009103; AAL42432.1; ALT_INIT.
CC EMBL; AE008068; AAK87218.1; -.
CC PIR; A97533; A97533.
CC HSSP; Q9NDH8; 10EP.
CC HAMAP; MF 00318; -.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; Enolase_C; 1.
CC Pfam; PF03952; Enolase_N; 1.
CC PRINTS; PR00148; ENOLASE.
CC ProDom; PD000902; Enolase; 1.
CC TIGRFAMs; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.
CC Complete proteome; Glycolysis; Lyase; Magnesium.
CC ACT_SITE 154 154 By similarity.
CC METAL 241 241 Magnesium (By similarity).
CC METAL 284 284 Magnesium (By similarity).
CC METAL 311 311 Magnesium (By similarity).
CC SEQUENCE 424 AA; 45142 MW; 84892005DEC0666F CRC64;
Query Match 65.8%; Score 1418; DB 1; Length 424;
Best Local Similarity 65.9%; Pred. No. 8.7e-76;
Matches 275; Conservative 59; Mismatches 81; Indels 2; Gaps 1;
QY 4 IMHVFARILDSRGNTPEAEVFLDDSGHVGAGVPSGASTGVHEAHELDDGGDRYLKGV 63
DB 4 ITDIAREILDSRGNTPEVDVYLEDGSMGAAVPSGASTGAHEAHELDDGGDRYLKGV 63
QY 64 LKAVENVEETDELAGLEADDORLIDAMIKLDTANKSRIGNAIIGVSMVAKAARD 123
DB 64 EKAVEAVNTEIPDAITGGFDAENQIQIDQMTALDGTNPKSRIGNAIIGVSLATAKAAAE 123
QY 124 SAGLPLFRYIGGNPAHLVPPVPMNIIINGGAHDSGVVDVQEFPIAGIETFSALRNGAE 183
DB 124 ASGLPELYRIGGNPAHLVPPVPMNIIINGGAHADNIDFQEFNLPVGAENIREAVRMGSE 183
QY 184 VYHALKSVIKEGLSTGLGDEGGFAPSVGSTREALDLIVEATEKAGTTPGKDIALALDVA 243
DB 184 VFHTLKKLSAQGHNTNVGDEGGFAPGLESAPALDFIMKSTIEKAGYRPGEDMYVGLDCA 243

QY 244 SSEFFKDTYHFEQGQHS--AAEMANVVAELVDAYPIVSIEDPLOEDDWEGYTNLTATIG 301
DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 303
244 STEFFKDGKYLVEGEGRTLEPGMAEYLAELVKNKPIIISVEDGMAEDWEGWKTITDLVG 303
QY 302 DKVQIVGDDFFVTNPERLKEGIKAANSILVKVNIQITLTETFDADVDMHAGYTSMS 361
DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 363
304 NKQVLGDDLFVTNSARLDDGKVMGANSILVKVNIQISLSETLDAVETAHKAGYTVMS 363
QY 362 HRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDAGVYAGRS 418
DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 420
364 HRSGETEDSTIADLAVATNCGQIKTGSLARSRLAKYNOLLRIEEMLMGPPQAAAYAGRS 420
RESULT 15
ENO_THETN STANDARD; PRT; 429 AA.
ID AC Q8R567;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocusNames=TTE1759;
OS Thermoanaerobacter tengcongensis.
OC Bacteria, Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OC NCBI_TaxID=119072;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=2192816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700 (2002).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE013129; AAM24953.1; -.
CC HSSP; P08324; 1E9I.
CC HAMAP; MF 00318; -.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; Enolase_C; 1.
CC Pfam; PF03952; Enolase_N; 1.
CC PRINTS; PR00148; ENOLASE.
CC ProDom; PD000902; Enolase; 1.
CC TIGRFAMs; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.
CC Complete proteome; Glycolysis; Lyase; Magnesium.
CC ACT_SITE 155 155 By similarity.
CC METAL 242 242 Magnesium (By similarity).
CC METAL 286 286 Magnesium (By similarity).
CC METAL 313 313 Magnesium (By similarity).
CC SEQUENCE 429 AA; 46296 MW; ABCC2445DEBA8D65 CRC64;
Query Match 65.5%; Score 1412; DB 1; Length 429;
Best Local Similarity 66.3%; Pred. No. 2e-75;

Matches		281;	Conservative	57;	Mismatches	82;	Indels	4;	Gaps	3;
Qy	1	VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDG-DRYL 59								
Db	1	MSSIIDYAREILDSRGNPTIEVEVELDSGAVGRAAVPSGASTGAFAEIELRDGDKSRYL 60								
Qy	60	GKGVLKAVENNEEIGDELAGLEADDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAK 119								
Db	61	GKGVLKAVQNVNDIIAPELIGMEAQDQVAIDKAMIELDGTENKSKLGANAILGVSLAVAK 120								
Qy	120	AAADSAGLPLFRYIGGNPAHVLVPMMNIINGGAHADSGVDVQEFMIAPIGAETFSALR 179								
Db	121	AAAEELGLPLYQYLGGVNAKTLVPMMNILNGGKHADNNVDIQEFMIMPVGAPNFKEARL 180								
Qy	180	NGAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALA 239								
Db	181	MCSEVYHSLKNVLHSLKGLSTTVGDEGGFAPNLTSEEAIKVILEAIEKAGVYVPGEDIVLA 240								
Qy	240	LDVASSBFFK-DGTYHFEKG--QHSAAEAMNVAELVDAYPIVSIEDPQEDDWEGYTNL 296								
Db	241	LDPAAATEMYKEDGKYHFEKGIVRTSEEMTEFWEQLVSKYPIVSIEDGLAEEDWNGKLL 300								
Qy	297	TATIGDKVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMVHRAGY 356								
Db	301	TERLGKKVQLVGDDLFVTNTERLSKGINMGVANSILIKLNQIGTTLTETDLDAIENAKRAGY 360								
Qy	357	TSMMSHRSGETETTIADLVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAG 416								
Db	361	TAIVSHRSGETETTIADLVVATNVGQIKTGAPARTDRVAKYNOLLRIEALGSVAOYPG 420								
Qy	417	RSAF 420								
Db	421	KNAF 424								

Search completed: September 29, 2005, 12:13:46
Job time : 92.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 20:45:32 ; Search time 4241.5 Seconds
(without alignments)
2568.796 Million cell updates/sec

Title: US-10-728-947-1

Perfect score: 1578

Sequence: 1 ggctgggagatgggtagtt.....ctcaagcagggaactgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1:	/cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/prodata/2/pubpna/US05_NEW_PUB.seq:*
4:	/cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
10:	/cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17:	/cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18:	/cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19:	/cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
20:	/cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21:	/cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22:	/cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
23:	/cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24:	/cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25:	/cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26:	/cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	9	US-09-860-768-1
2	1578	100.0	1578	20	US-10-728-947-1
3	1578	100.0	3309400	9	US-09-738-626-1
4	1576.4	99.9	1578	20	US-09-860-768-3
5	1576.4	99.9	1578	9	US-10-728-947-3
6	1403.4	88.9	1405	21	US-10-494-836-7
7	1396.4	88.5	1398	19	US-10-781-014-71

8	1275	80.8	1275	9	US-09-738-626-1085
9	908.6	57.6	1275	17	US-10-282-122A-17724
10	698.4	44.3	9025608	15	US-10-156-761-1
11	897	44.2	1284	15	US-10-156-761-3520
12	679.8	43.1	1269	17	US-10-369-493-32027
13	647	41.0	1287	17	US-10-282-122A-25949
14	632.6	40.1	45191	15	US-10-080-170-649
15	632.6	40.1	45191	19	US-10-080-170-649
16	632.6	40.1	45191	19	US-10-468-356-649
17	632.4	40.1	1317	17	US-10-282-122A-26512
18	626.2	39.7	1290	17	US-10-282-122A-28306
19	584.6	37.0	2256646	19	US-10-470-565-1
20	583.4	37.0	1344	17	US-10-282-122A-27498
21	582	36.9	1269	17	US-10-369-493-34244
22	575.4	36.5	1269	17	US-10-369-493-35383
23	574	36.4	1275	17	US-10-369-493-35506
24	574	36.4	1281	17	US-10-369-493-40598
25	569.8	36.1	1278	17	US-10-369-493-41573
26	569	36.1	1272	17	US-10-369-493-44349
27	567.4	36.0	1260	17	US-10-369-493-38436
28	567.4	36.0	1269	17	US-10-369-493-38909
29	561.8	35.6	1272	17	US-10-369-493-31449
30	557	35.3	1278	17	US-10-369-493-45417
31	552	35.0	1290	17	US-10-369-493-43178
32	542	34.3	1281	17	US-10-282-122A-11620
33	531.8	33.7	1290	9	US-09-815-242-7834
34	531.8	33.7	1290	17	US-10-282-122A-30326
35	530.8	33.6	1281	17	US-10-369-493-44693
36	528	33.5	1287	17	US-10-369-493-32181
37	522.8	33.1	1278	17	US-10-369-493-28375
38	522.8	33.1	1281	17	US-10-369-493-31134
39	522.8	33.1	1281	17	US-10-282-122A-13098
40	515.2	32.6	1404	17	US-10-282-122A-33598
41	514.8	32.6	1269	17	US-10-369-493-47120
42	513.6	32.5	1293	17	US-10-369-493-46986
43	509.4	32.3	1293	17	US-10-369-493-37536
44	507.6	32.2	1293	9	US-09-974-300-664
45	504.8	32.0	1293	17	US-10-282-122A-9983

ALIGNMENTS

RESULT 1

US-09-860-768-1
; Sequence 1, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bertina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860.768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-1

Query Match 100.0%; Score 1578; DB 9; Length 1578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTGGGATATGGGTAGTTTCGCCACTAATTTCACTGATTCCTTCGCAACAGAGA 60
Db 1 GGCTGGGATATGGGTAGTTTCGCCACTAATTTCACTGATTCCTTCGCAACAGAGA 60

QY 61 TTCTGTCACCAATTCGGGTAGACGTGATTGAAGACATTTGATCAGGTGAATAATTCTAG 120
 Db 61 TTCTGTCACCAATTCGGGTAGACGTGATTGAAGACATTTGATCAGGTGAATAATTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCATAGAGGCGCCACAGTGGCTGAAATCATGCAAGTATTTCGCTGCG 180
 Db 121 TTAGCTCCCAAGTTGGCATAGAGGCGCCACAGTGGCTGAAATCATGCAAGTATTTCGCTGCG 180
 QY 181 GAAATTCGATCTCCGGGCTTAACCCAAACCGTTCGAGGCGAGAGTTTCTGATGACGGT 240
 Db 181 GAAATTCGATCTCCGGGCTTAACCCAAACCGTTCGAGGCGAGAGTTTCTGATGACGGT 240
 QY 241 TCCACGGTTCGAGGCTTCATCCGCGGATCCACGGGCTCCAGAGGCTCATGAG 300
 Db 241 TCCACGGTTCGAGGCTTCATCCGCGGATCCACGGGCTCCAGAGGCTCATGAG 300
 QY 301 CTGCGTGAAGGCTGCGATCGCTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 360
 Db 301 CTGCGTGAAGGCTGCGATCGCTACCTGGGCAAGGCGTTTGAAGGCGAGTTGMAAACGTC 360
 QY 361 AACGAAGAAATCGGCGAGCTCGCTAGCTGAGGCTGACGATCAGCGCTCATCGAC 420
 Db 361 AACGAAGAAATCGGCGAGCTCGCTAGCTGAGGCTGACGATCAGCGCTCATCGAC 420
 QY 421 GAAGCAATGATCAAGCTTGAAGGCTCGGCAACCAAGTCCGCGCTGAGGCTGCAACGCAATC 480
 Db 421 GAAGCAATGATCAAGCTTGAAGGCTCGGCAACCAAGTCCGCGCTGAGGCTGCAACGCAATC 480
 QY 481 CTGCGTTCATCGCTGTTGCAAGGCTGCTGATTCGCGAGGCTCCCACTGCTTC 540
 Db 481 CTGCGTTCATCGCTGTTGCAAGGCTGCTGATTCGCGAGGCTCCCACTGCTTC 540
 QY 541 CGCTACATCGGTGGAACCAACGACACAGTCTTCCAGTTTCCAAATGATGAACATCATCAAC 600
 Db 541 CGCTACATCGGTGGAACCAACGACACAGTCTTCCAGTTTCCAAATGATGAACATCATCAAC 600
 QY 601 GGTGCGCTCAGCTGATCTCGGCTGTTGAGTTCAGGAAATCATGATCGCTCCAAATCGGT 660
 Db 601 GGTGCGCTCAGCTGATCTCGGCTGTTGAGTTCAGGAAATCATGATCGCTCCAAATCGGT 660
 QY 661 GCAGAGACCTCTCTGAGGCTCTCGCAACGCGGCGAGGCTTACACGCACTGAAGTCC 720
 Db 661 GCAGAGACCTCTCTGAGGCTCTCGCAACGCGGCGAGGCTTACACGCACTGAAGTCC 720
 QY 721 GTTCATCAAGAAAGGCGCTCTCCAGCACTTGGCGATGAGGGGCTTCGCTCTCTCC 780
 Db 721 GTTCATCAAGAAAGGCGCTCTCCAGCACTTGGCGATGAGGGGCTTCGCTCTCTCC 780
 QY 781 GTGCGCTCACCCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTC 840
 Db 781 GTGCGCTCACCCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTC 840
 QY 841 ACCCCAGGCAAGGACATCGCTCTGCTCTGAGGCTTCTCTCTGAGTCTTCAAGGAC 900
 Db 841 ACCCCAGGCAAGGACATCGCTCTGCTCTGAGGCTTCTCTCTGAGTCTTCAAGGAC 900
 QY 901 GGCACTTACCACTTCCAGAGTGGCGACGACTCCGCACTGAGTGAAGGCAACGTTTACGCT 960
 Db 901 GGCACTTACCACTTCCAGAGTGGCGACGACTCCGCACTGAGTGAAGGCAACGTTTACGCT 960
 QY 961 GAGCTGTTGACGCTGATCCCAATCGTCTCATGAGGACCCACTGCAAGAGATGACTGG 1020
 Db 961 GAGCTGTTGACGCTGATCCCAATCGTCTCATGAGGACCCACTGCAAGAGATGACTGG 1020
 QY 1021 GAGGTTACCAACCTCAGGCGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAAGTCC 1080
 Db 1021 GAGGTTACCAACCTCAGGCGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAAGTCC 1080
 QY 1081 TTCTTCGTCAACCAACCTCAGGCGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAAGTCC 1140
 Db 1081 TTCTTCGTCAACCAACCTCAGGCGCTGAGGAGGCGATCGCTTAAGAGGCTGCCAAGTCC 1140

QY 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTCCACGAGACCTTCGACGCTGTCCGACATG 1200
 Db 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTCCACGAGACCTTCGACGCTGTCCGACATG 1200
 QY 1201 GCTCACCGCGAGGCTACCTCCATGATGTCACCGTTCGGTTCGGTGAAGCCGAGGACACC 1260
 Db 1201 GCTCACCGCGAGGCTACCTCCATGATGTCACCGTTCGGTTCGGTGAAGCCGAGGACACC 1260
 QY 1261 ACCATTCCTGACCTCGCAGTTGCACTCACTGTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
 Db 1261 ACCATTCCTGACCTCGCAGTTGCACTCACTGTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
 QY 1321 CGTTCGACCGCTGCGCAAAAGTACCAACAGCTTCTCCGCACTCGAGAGCTGCTTGGGAC 1380
 Db 1321 CGTTCGACCGCTGCGCAAAAGTACCAACAGCTTCTCCGCACTCGAGAGCTGCTTGGGAC 1380
 QY 1381 GCGCGCTCTACGAGGTCGACGCAATCCAGCTTTCCAGGCTTAATAAAGGCGCTT 1440
 Db 1381 GCGCGCTCTACGAGGTCGACGCAATCCAGCTTTCCAGGCTTAATAAAGGCGCTT 1440
 QY 1441 TTGACGCGCGGTAACCTCAAGTTGCGGGCGCTGCTTACTACTGTTACTGCTGT 1500
 Db 1441 TTGACGCGCGGTAACCTCAAGTTGCGGGCGCTGCTTACTACTGTTACTGCTGT 1500
 QY 1501 GACTATGATCGAGGATTATGCAAGGAGCAAGAACTCATAAAGGCTTGTCTCTGCT 1560
 Db 1501 GACTATGATCGAGGATTATGCAAGGAGCAAGAACTCATAAAGGCTTGTCTCTGCT 1560
 QY 1561 CAAGCAGGGAACGCTT 1578
 Db 1561 CAAGCAGGGAACGCTT 1578

RESULT 2
 US-10-728-947-1
 ; Sequence 1 Application US/10728947
 ; Publication No. US20040220394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jörn
 ; APPLICANT: Bache, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10728,947
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1578
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; US-10-728-947-1

Query Match 100.0%; Score 1578; DB 20; Length 1578;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTGGGGAATGAGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAACAAGA 60
 Db 1 GGCTGGGGAATGAGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAACAAGA 60
 QY 61 TTCTGCAACAATTTGGGTGATGATGATGAACACATTTGATCACTGATTAATTTCTAG 120
 Db 61 TTCTGCAACAATTTGGGTGATGATGATGAACACATTTGATCACTGATTAATTTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCATAGAGGCGCCACAGTGGCTGAAATCATGCAAGTATTTCGCTGCG 180
 Db 121 TTAGCTCCCAAGTTGGCATAGAGGCGCCACAGTGGCTGAAATCATGCAAGTATTTCGCTGCG 180

181 GAAATTTCTGACTCCCGCGGTACCAACCGTTCGAGCGAGGTTTCTTGGATGACGGT 240
 Db GAAATTTCTGACTCCCGCGGTACCAACCGTTCGAGCGAGGTTTCTTGGATGACGGT 240
 241 TCCACCGGTGTCGACGGTGTTCATCCGCGGATCCACCGGCGTCCACGAGGCTCATGAG 300
 Db TCCACCGGTGTCGACGGTGTTCATCCGCGGATCCACCGGCGTCCACGAGGCTCATGAG 300
 301 CTGCGTGAAGGTGGCGATCGCTACCTTGGCGAAGGGCGTTTGAAGCAGTTGAAGCGTC 360
 Db CTGCGTGAAGGTGGCGATCGCTACCTTGGCGAAGGGCGTTTGAAGCAGTTGAAGCGTC 360
 361 AACGAGAAATCCGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTTCATCGAC 420
 Db AACGAGAAATCCGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTTCATCGAC 420
 421 GAAGCAATGATCAAGCTTGAATGSCACCGCAACCAAGTCCCGCTGGGTGCAACGCAATC 480
 Db GAAGCAATGATCAAGCTTGAATGSCACCGCAACCAAGTCCCGCTGGGTGCAACGCAATC 480
 481 CTTGGTGTTCATGAGCTGTGCAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTTTC 540
 Db CTTGGTGTTCATGAGCTGTGCAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTTTC 540
 541 CGCTACATCGGTGGACCAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
 Db CGCTACATCGGTGGACCAACGACACGCTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
 601 GGTGGCGCTCAGCTGACTCCGCTGTGAGCTTTCAGGAAATTCATGATCGCTCCAACTCGT 660
 Db GGTGGCGCTCAGCTGACTCCGCTGTGAGCTTTCAGGAAATTCATGATCGCTCCAACTCGT 660
 661 GCAGAGACTTCTCTGAGGCTTCCGAAACGCGGCGAGGCTTACCAAGCTGAAGTCC 720
 Db GCAGAGACTTCTCTGAGGCTTCCGAAACGCGGCGAGGCTTACCAAGCTGAAGTCC 720
 721 GTCATCAAGGAAAGGCGCTTCCAGCGGCTTGGGATGAGGCGGCTTCCCTCTCC 780
 Db GTCATCAAGGAAAGGCGCTTCCAGCGGCTTGGGATGAGGCGGCTTCCCTCTCC 780
 781 GTGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATCAGAGGCTGCTTC 840
 Db GTGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATCAGAGGCTGCTTC 840
 841 ACCCAGGCAAGGACATCGCTTGTCTGAGGCTTCTGAGGCTTCTGAGTCTTCAAGGAC 900
 Db ACCCAGGCAAGGACATCGCTTGTCTGAGGCTTCTGAGGCTTCTGAGTCTTCAAGGAC 900
 901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCGGAGCTGAGATGCGCAACGTTTACGCT 960
 Db GGCACCTACCACTTCGAAGGTGGCGAGCACTCGGAGCTGAGATGCGCAACGTTTACGCT 960
 961 GAGCTGTTGAGCGGTACCAATCGTCTCCATCGAGGACCACTGCGAGGATGACTGG 1020
 Db GAGCTGTTGAGCGGTACCAATCGTCTCCATCGAGGACCACTGCGAGGATGACTGG 1020
 1021 GAGGTTACCAACCTTCAACGCAACCATCGGCGAAGGTTTCAGATCGTTGGCGAGAC 1080
 Db GAGGTTACCAACCTTCAACGCAACCATCGGCGAAGGTTTCAGATCGTTGGCGAGAC 1080
 1081 TTCTTGTTCACCAACCTTGAAGGCTTGAAGGAGGCGATCGTTAAGAGGCTTCCAACTCC 1140
 Db TTCTTGTTCACCAACCTTGAAGGCTTGAAGGAGGCGATCGTTAAGAGGCTTCCAACTCC 1140
 1141 ATCTCTGTTAAGGTGAACAGATCGGTACCTTCCAGGAGGCTTGAAGGCTTGGCATG 1200
 Db ATCTCTGTTAAGGTGAACAGATCGGTACCTTCCAGGAGGCTTGAAGGCTTGGCATG 1200
 1201 GCTCAGCGGCGAGGCTTACACCTTCCATGATGTCCACCGTTCGCGTGAAGGAGGACCC 1260
 Db GCTCAGCGGCGAGGCTTACACCTTCCATGATGTCCACCGTTCGCGTGAAGGAGGACCC 1260

1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA 1320
 Db ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA 1320
 1321 CGTTCCGACCGGTGTCGCAAAAGTACAAACGAGTTTCCCGATCGAGCAGCTGTTGGCGAC 1380
 Db CGTTCCGACCGGTGTCGCAAAAGTACAAACGAGTTTCCCGATCGAGCAGCTGTTGGCGAC 1380
 1381 GCCGCGCTACGAGGTCGAGCGCATTCACCGCTTCCAGGCTTAATAAAGCGCTT 1440
 Db GCCGCGCTACGAGGTCGAGCGCATTCACCGCTTCCAGGCTTAATAAAGCGCTT 1440
 1441 TTCCAGCGCCGCTAACCTCAAGGTTGCGGCGCTGTTGCTTACTTACTTCTGCTT 1500
 Db TTCCAGCGCCGCTAACCTCAAGGTTGCGGCGCTGTTGCTTACTTACTTCTGCTT 1500
 1501 GACTATGATCGAGGATTAAGCAAGAGCAAGAACTCATAAAGGCTTGTCTGCTT 1560
 Db GACTATGATCGAGGATTAAGCAAGAGCAAGAACTCATAAAGGCTTGTCTGCTT 1560
 1561 CAAGCGGGAACGCTT 1578
 Db CAAGCGGGAACGCTT 1578

RESULT 3

US-09-738-626-1
 ; Sequence 1, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: JP 99/37484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patent in ver. 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3309400
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-1

Query Match 100.0%; Score 1578; DB 9; Length 3309400;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GCCTGGGATATGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 60
 Db 1034799 GCCTGGGATATGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 1034858
 61 TTCTGTCAACAAATGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 120
 Db 1034859 TTCTGTCAACAAATGGGTAGTGTTCGCACTAATTTCAACTGATGCTCATCGAAACAAGA 1034918
 121 TTAGCTCCCAAGTGGCATAGGAGGCGACAGTGGCTGAAATCATGACGATATTCGCTCGC 180
 Db 1034919 TTAGCTCCCAAGTGGCATAGGAGGCGACAGTGGCTGAAATCATGACGATATTCGCTCGC 1034978

No patent data - 1999-11-16
 provided

QY 181 GAAATTCGACTCCCGCGGTAAACCAACCGTCGAGGAGAGGTTTTCTTGATGACGGT 240
Db 1034979 GAAATTCGACTCCCGCGGTAAACCAACCGTCGAGGAGAGGTTTTCTTGATGACGGT 1035038
QY 241 TCCCAAGGTGTCGAGGTGTTCCATCCGCGCGATCCACGCGGTCCAGAGGCTCATGAG 300
Db 1035039 TCCCAAGGTGTCGAGGTGTTCCATCCGCGCGATCCACGCGGTCCAGAGGCTCATGAG 1035098
QY 301 CTGCGTGAACGCGGTGCGATCGCTACCTTGGGCAAGGCGGTTTTGAAGGCAAGTTGMAAACGTC 360
Db 1035099 CTGCGTGAACGCGGTGCGATCGCTACCTTGGGCAAGGCGGTTTTGAAGGCAAGTTGMAAACGTC 1035158
QY 361 AACGAAGAAATCGGCGAAGAGCTCGCTGCGCTAGAGGCTGACGATCAACGCGCTCATGAC 420
Db 1035159 AACGAAGAAATCGGCGAAGAGCTCGCTGCGCTAGAGGCTGACGATCAACGCGCTCATGAC 1035218
QY 421 GAAGCAATGATCAAGCTTGATGGCAACCGCAACCAAGTCCCGCTGGGTGCAACGCAATC 480
Db 1035219 GAAGCAATGATCAAGCTTGATGGCAACCGCAACCAAGTCCCGCTGGGTGCAACGCAATC 1035278
QY 481 CTGCGTGAACGCGGTGCGATCGCTACCTTGGGCAAGGCGGTTTTGAAGGCAAGTTGMAAACGTC 540
Db 1035279 CTGCGTGAACGCGGTGCGATCGCTACCTTGGGCAAGGCGGTTTTGAAGGCAAGTTGMAAACGTC 1035338
QY 541 CGCTACATCGGTGGAACCAACGCAACAGTTTCTTCCAGTTCCAAATGATGAACATCATCAAC 600
Db 1035339 CGCTACATCGGTGGAACCAACGCAACAGTTTCTTCCAGTTCCAAATGATGAACATCATCAAC 1035398
QY 601 GGTGGCGCTCAGCTGATCCGCGTGTGAGCTTTCAGGAAATTCATGATCGCTCCATCGGT 660
Db 1035399 GGTGGCGCTCAGCTGATCCGCGTGTGAGCTTTCAGGAAATTCATGATCGCTCCATCGGT 1035458
QY 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACACGCACTGAAGTCC 720
Db 1035459 GCAGAGACTTCTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACACGCACTGAAGTCC 1035518
QY 721 GTCTACAAAGGAAAGGCGCTGTCACCGCACTTGGCGATGAGGCGGCTTCGCTCCCTCC 780
Db 1035519 GTCTACAAAGGAAAGGCGCTGTCACCGCACTTGGCGATGAGGCGGCTTCGCTCCCTCC 1035578
QY 781 GTCGCTCCACCGGTGAGGCTTTCGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
Db 1035579 GTCGCTCCACCGGTGAGGCTTTCGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 1035638
QY 841 ACCCGAGCAAGGACATCGCTCTTCTGCTGAGAGCTTCTCTCTGAGTCTTCTCAAGGAC 900
Db 1035639 ACCCGAGCAAGGACATCGCTCTTCTGCTGAGAGCTTCTCTCTGAGTCTTCTCAAGGAC 1035698
QY 901 GGCACTTACCACTTCGAAAGGTGGCGAGCACTCCGCGAGTGAATGGCAAAAGCTTACGCT 960
Db 1035699 GGCACTTACCACTTCGAAAGGTGGCGAGCACTCCGCGAGTGAATGGCAAAAGCTTACGCT 1035758
QY 961 GAGCTCGTTGAGCGGTACCCATCGCTCTCCATCGAGGACCCACTGCGAGGAGATGACTGG 1020
Db 1035759 GAGCTCGTTGAGCGGTACCCATCGCTCTCCATCGAGGACCCACTGCGAGGAGATGACTGG 1035818
QY 1021 GAGGTTTACCAACCTCAGCGCAACCATCGGCGACAGGTTTCAGATCGTTGGCGAGCAG 1080
Db 1035819 GAGGTTTACCAACCTCAGCGCAACCATCGGCGACAGGTTTCAGATCGTTGGCGAGCAG 1035878
QY 1081 TTCTTCGTCACCAACCTTCAGCGCTGAAAGGAGGCACTGCTTAAAGAGGCTGCCAATCC 1140
Db 1035879 TTCTTCGTCACCAACCTTCAGCGCTGAAAGGAGGCACTGCTTAAAGAGGCTGCCAATCC 1035938
QY 1141 ATCTCGTTTAAAGTGAACCAATCGGTACCTTCAACGAGCTTCGAGCGCTTCGACATG 1200
Db 1035939 ATCTCGTTTAAAGTGAACCAATCGGTACCTTCAACGAGCTTCGAGCGCTTCGACATG 1035998
QY 1201 GCTCACCGCGAGGCTTACACCTTCATGATGTCCTCCACCGTTCGCGTGAACGAGGACACC 1260
Db 1035999 GCTCACCGCGAGGCTTACACCTTCATGATGTCCTCCACCGTTCGCGTGAACGAGGACACC 1036058

QY 1261 ACCATTCGACTCCGCGGTGCACTCAACTGTGGCCAGATCAAGACTGTGTCTCCAGCA 1320
Db 1036059 ACCATTCGACTCCGCGGTGCACTCAACTGTGGCCAGATCAAGACTGTGTCTCCAGCA 1036118
QY 1321 CGTTCGAGCGGTGTCGCAAGGTACCAACGAGCTTCTCGCATCGAGGAGCTGCTTGGCGAC 1380
Db 1036119 CGTTCGAGCGGTGTCGCAAGGTACCAACGAGCTTCTCGCATCGAGGAGCTGCTTGGCGAC 1036178
QY 1381 GCGCGCGTCTACGCGAGTCCGAGCGATCCACGCTTTCAGGCTTAAATAAAGCGCTT 1440
Db 1036179 GCGCGCGTCTACGCGAGTCCGAGCGATCCACGCTTTCAGGCTTAAATAAAGCGCTT 1036238
QY 1441 TTGCAAGCGCCGCTTAACTCAAGGTTGCGCGGCTGCTGCTTACTACTGTACTGTGT 1500
Db 1036239 TTGCAAGCGCCGCTTAACTCAAGGTTGCGCGGCTGCTGCTTACTACTGTACTGTGT 1036298
QY 1501 GACTATGATCGAGGATATGGCAAGGAGGAGGAACTCATAAAGGCTTGTCTCTGCT 1560
Db 1036299 GACTATGATCGAGGATATGGCAAGGAGGAGGAACTCATAAAGGCTTGTCTCTGCT 1036358
QY 1561 CAAGCAGGGAACGCTGCTT 1578
Db 1036359 CAAGCAGGGAACGCTGCTT 1036376

RESULT 4

US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match 99.9%; Score 1576.4; DB 9; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGGGGATATGGGTAGTGTTCGCCACTAAATTTCAACTGATTCCTCATCGAAACAAGA 60
Db 1 GGCTGGGGATATGGGTAGTGTTCGCCACTAAATTTCAACTGATTCCTCATCGAAACAAGA 60
QY 61 TTCTGTCACAACTATGGGTGTAGACGTGATTAAGAGACATTTGATCACTGAATTAATCTAG 120
Db 61 TTCTGTCACAACTATGGGTGTAGACGTGATTAAGAGACATTTGATCACTGAATTAATCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTAATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTAATTCGCTCGC 180
QY 181 GAAATTCCTGACTCCCGCGGTAAACCCAAACCGTCCGAGGAGAGGTTTTCTCGGATGACGGT 240
Db 181 GAAATTCCTGACTCCCGCGGTAAACCCAAACCGTCCGAGGAGAGGTTTTCTCGGATGACGGT 240
QY 241 TCCCAAGGCTGTGCGAGGTTTCATCCGCGCATCCACCGGCGCTCCACGAGGCTCATGAG 300
Db 241 TCCCAAGGCTGTGCGAGGTTTCATCCGCGCATCCACCGGCGCTCCACGAGGCTCATGAG 300
QY 301 CTGCGTGAACGCTGCGATCGCTTACCTGGGCGAAGGCGGTTTTTGAAGGAGGATGTAACGTC 360

301 CTGCGTGA CGGTGGCGATCGTACCTGGGCAAGGGCGTTTGAAGCGAGTTGAAAAGCTC 360
361 AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGGGCTCATCGAC 420
361 AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGGGCTCATCGAC 420
421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTCGGGTGCMAACGCAATC 480
421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTCGGGTGCMAACGCAATC 480
481 CTTGGTGTTCATGCTGTGTCGAAAGCTGTGCTGATTCGCGACGGCTCCCACTGTTTC 540
481 CTTGGTGTTCATGCTGTGTCGAAAGCTGTGCTGATTCGCGACGGCTCCCACTGTTTC 540
541 CGCTACATCGGTGGGCAACGACGCTTCTCCAGTTCCTCCAGTTCCTCCATCATCAAC 600
541 CGCTACATCGGTGGGCAACGACGCTTCTCCAGTTCCTCCAGTTCCTCCATCATCAAC 600
601 GGTGGCGCTCACGCTGACTCCGGTGTGAGCTTTCAGGAATTCATGATCGTCCCAATCGGT 660
601 GGTGGCGCTCACGCTGACTCCGGTGTGAGCTTTCAGGAATTCATGATCGTCCCAATCGGT 660
661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGCTTACCAACGCTGAAATGCC 720
661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGCTTACCAACGCTGAAATGCC 720
721 GTCATCAAGAAAGGGCTCTCCAGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
721 GTCATCAAGAAAGGGCTCTCCAGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
781 GTCCGCTCCACCGCTGAGGCTCTGAGCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
781 GTCCGCTCCACCGCTGAGGCTCTGAGCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
841 ACCCCAGGCAAGGACATCGCTCTGCTCTGGAGTTCCTCTGAGTTCCTCAAGGAC 900
841 ACCCCAGGCAAGGACATCGCTCTGCTCTGGAGTTCCTCTGAGTTCCTCAAGGAC 900
901 GGCACCTACCTTCGAAAGTGGCCAGCACTCGCGAGCTGAGATGCAAGCTTACGCT 960
901 GGCACCTACCTTCGAAAGTGGCCAGCACTCGCGAGCTGAGATGCAAGCTTACGCT 960
961 GAGCTGTGAGCGCTTACCAATCGTCTCCATCGAGGACCCACTCGAGGAATGACTGG 1020
961 GAGCTGTGAGCGCTTACCAATCGTCTCCATCGAGGACCCACTCGAGGAATGACTGG 1020
1021 GAGGGTTACACCAACCTCACCGCAACCATCGGCGACAAAGTTTCAGATCGTGGCGACGAC 1080
1021 GAGGGTTACACCAACCTCACCGCAACCATCGGCGACAAAGTTTCAGATCGTGGCGACGAC 1080
1081 TTCTTCGTCAACCAACCTGAGGCTGAAAGGAGGATCGCTAAGAGGCTGCCAATCC 1140
1081 TTCTTCGTCAACCAACCTGAGGCTGAAAGGAGGATCGCTAAGAGGCTGCCAATCC 1140
1141 ATCTCTGGTTAAGGTGAACAGATCGTACCTTACCGGACCTTTCGACGCTGTGCACATG 1200
1141 ATCTCTGGTTAAGGTGAACAGATCGTACCTTACCGGACCTTTCGACGCTGTGCACATG 1200
1201 GCTCACCGCGAGGCTTACACCTTCCATGATGTCCACCGTTCGGTGGAGACCGGACAC 1260
1201 GCTCACCGCGAGGCTTACACCTTCCATGATGTCCACCGTTCGGTGGAGACCGGACAC 1260
1261 ACCATTGCTGACCTCGAGTTGACCTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
1261 ACCATTGCTGACCTCGAGTTGACCTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
1321 CGTTCGACCGGTGTCGAAAGTACAAAGCTTCTCCGATCGAGCAGCTGTGTGGCGAC 1380
1321 CGTTCGACCGGTGTCGAAAGTACAAAGCTTCTCCGATCGAGCAGCTGTGTGGCGAC 1380
1381 GCGGGGCTTACGCGAGTGGCGCAATTCACCGCTTTCAGGGCTTAATAAAGGCTT 1440

1381 GCGGGGCTTACGCGAGTGGCGCAATTCACCGCTTTCAGGGCTAAATAAAGGCTT 1440
1441 TTCGACGCGCGTAACTCAAGCTTTCGCGGCGTTCGCTTACTACTGTTACTGGTCT 1500
1441 TTCGACGCGCGTAACTCAAGCTTTCGCGGCGTTCGCTTACTACTGTTACTGGTCT 1500
1501 GACTATGATCGAGGATTTATGCAAGCAAGAAACTCATAAAGGCTTGTTCCTGTCT 1560
1501 GACTATGATCGAGGATTTATGCAAGCAAGAAACTCATAAAGGCTTGTTCCTGTCT 1560
1561 CAAGCAGGGAACGTGCTT 1578
1561 CAAGCAGGGAACGTGCTT 1578

RESULT 5
US-10-728-947-3
; Sequence 3, Application US/10728947
; Publication No. US20040220394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10/728,947
; CURRENT FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-3

Query Match 99.9%; Score 1576.4; DB 20; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCGTGGGATATGGTAGTCTTTCGCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
DB 1 GCGTGGGATATGGTAGTCTTTCGCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
QY 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTCGAAGACATTTGATCAGCTGAATATCTAG 120
DB 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTCGAAGACATTTGATCAGCTGAATATCTAG 120
QY 121 TTAGCTCCCAAGTTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGATATTCCTCGC 180
DB 121 TTAGCTCCCAAGTTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGATATTCCTCGC 180
QY 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGGTTTTCTGATGACCGGT 240
DB 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGGTTTTCTGATGACCGGT 240
QY 241 TCCCAAGGTGTGCGAGGTTCCTCATCCGCGCATCCACCGGCTCCAGAGGCTCATGAG 300
DB 241 TCCCAAGGTGTGCGAGGTTCCTCATCCGCGCATCCACCGGCTCCAGAGGCTCATGAG 300
QY 301 CTGCGTGAACGCTGGCGATCGCTACTCTGGGCAAGGCGGTTTTGAAGGAGTTGAAAACGTC 360
DB 301 CTGCGTGAACGCTGGCGATCGCTACTCTGGGCAAGGCGGTTTTGAAGGAGTTGAAAACGTC 360
QY 361 AACGAGAAATCGGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGAC 420
DB 361 AACGAGAAATCGGCGACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTTCCTCCAGTTCCTCCATCATCAAC 480

Db 421 GAAGCAATGATCAAGCTTGATGGCAACCGCAACAAGTCCCGCTGGGTGCAACGCAATC 480
Qy 481 CTGTGTGTTTCATGGCTGTTGCAAAAGCTGCTGCTGATTCGGCAGGCTCCCACTGTTT 540
Db 481 CTGTGTGTTTCATGGCTGTTGCAAAAGCTGCTGCTGATTCGGCAGGCTCCCACTGTTT 540
Qy 541 CGCTACATCGGTGGACCAACGCAACAGTCTTTCAGTTTCCATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAACGCAACAGTCTTTCAGTTTCCATGATGAACATCATCAAC 600
Qy 601 GGTGGCGCTCAGCGTACTCGGGTCTTCAGCTTTCAGGAATTCATGATCGTCCCAATCGGT 660
Db 601 GGTGGCGCTCAGCGTACTCGGGTCTTCAGCTTTCAGGAATTCATGATCGTCCCAATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGGAGTCTACACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGGAGTCTACACGCACTGAAGTCC 720
Qy 721 GTCATCAAGGAAAGGGCTGTGCAACGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGGAAAGGGCTGTGCAACGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
Qy 781 GTCGGCTCCACCGTGAGGCTCTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
Db 781 GTCGGCTCCACCGTGAGGCTCTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
Qy 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTCTCTGAGTTCTTCAAGGAC 900
Qy 901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGGCAGCTGAGATGGCAAAAGTTACGCT 960
Db 901 GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGGCAGCTGAGATGGCAAAAGTTACGCT 960
Qy 961 GAGCTCGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGGAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGGAGATGACTGG 1020
Qy 1021 GAGGTTTACCAACCTCACCGCAACCATCGGCGACAGGTTTCAGATCGTTGGCGAGAC 1080
Db 1021 GAGGTTTACCAACCTCACCGCAACCATCGGCGACAGGTTTCAGATCGTTGGCGAGAC 1080
Qy 1081 TTCTTGTGTCACCAACCTGAGCGCTGAGAGGGCATCGCTAAGAGGCTGGCAACTCC 1140
Db 1081 TTCTTGTGTCACCAACCTGAGCGCTGAGAGGGCATCGCTAAGAGGCTGGCAACTCC 1140
Qy 1141 ATCTGTTTAAAGTGAACAGATCGGTACCTCCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db 1141 ATCTGTTTAAAGTGAACAGATCGGTACCTCCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Qy 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCTCACCGGTCGGGTGAGACCGAGGACAC 1260
Db 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCTCACCGGTCGGGTGAGACCGAGGACAC 1260
Qy 1261 ACCATGCTGACCTCGCAGTTGCACTCAACTGTTGCGCAGATCAAGACTGGTGTCCAGCA 1320
Db 1261 ACCATGCTGACCTCGCAGTTGCACTCAACTGTTGCGCAGATCAAGACTGGTGTCCAGCA 1320
Qy 1321 GGTTCGAGCGGTGCGCAAGTACCAACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC 1380
Db 1321 GGTTCGAGCGGTGCGCAAGTACCAACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCCGGCTCTACGAGGTGCGAGGCAATCCACGCTTTCAGGGCTAAATAAAGCGCTT 1440
Db 1381 GCCGGCTCTACGAGGTGCGAGGCAATCCACGCTTTCAGGGCTAAATAAAGCGCTT 1440
Qy 1441 TTGACCGCGGTAACTCAAGGTGCGGGCGTGTGTCCTTACTACTGTTACTGTTGT 1500
Db 1441 TTGACCGCGGTAACTCAAGGTGCGGGCGTGTGTCCTTACTACTGTTACTGTTGT 1500
Qy 1501 GACTATGATCGAGGATTAAGCAAGCAGAGAAAGTCAATAAAGGCTGCTGCTGCTCT 1560

Db 1501 GACTATGATCGAGGATTAAGCAAGCAGAGAAAGTCAATAAAGGCTGCTGCTGCTCT 1560
Qy 1561 CAAGCAGGAAACGTGCTT 1578
Db 1561 CAAGCAGGAAACGTGCTT 1578
RESULT 6
US-10-494-836-7
; Sequence 7, Application US/10494836
; Publication No. US20050014233A1
; GENERAL INFORMATION:
; APPLICANT: Zeider, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Burkhard
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopprogge, Corinna
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy produc
; FILE REFERENCE: BGI-167US
; CURRENT APPLICATION NUMBER: US/10/494,836
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: PCT/EP02/12135
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 101 54 270.4
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 7
; LENGTH: 1405
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1375)
; OTHER INFORMATION: RXA00235
US-10-494-836-7

Query Match 88.9%; Score 1403.4; DB 21; Length 1405;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 51 CGAAACAAGATTCTGTCACAAATTTGGGTGTAGACGTGATTTGAAGACATTTGATCACGTGA 110
Db 1 CGAAACAAGATTCTGTCACAAATTTGGGTGTAGACGTGATTTGAAGACATTTGATCACGTGA 60
Qy 111 ATAATTCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGCAAAATCATGCAAGT 170
Db 61 ATAATTCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGCAAAATCATGCAAGT 120
Qy 171 ATTCTGCTCGCAAAATTTCTGACCTCCCGCGGTAAACCCAAACCGTTCGAGCAGAGGTTTTCTT 230
Db 121 ATTCTGCTCGCAAAATTTCTGACCTCCCGCGGTAAACCCAAACCGTTCGAGCAGAGGTTTTCTT 180
Qy 231 GGATGACGTTTCCACGCGTGTCCAGGTTTCCATCGGGGCAATCCACCGGCGTCCACGA 290
Db 181 GGATGACGTTTCCACGCGTGTCCAGGTTTCCATCGGGGCAATCCACCGGCGTCCACGA 240
Qy 291 GGTCTCATGAGCTCGGTGACGCGTGGCGATCGTACCTGGGCAAGGGCGGTTTTGAAGGCGAGT 350
Db 241 GGTCTCATGAGCTCGGTGACGCGTGGCGATCGTACCTGGGCAAGGGCGGTTTTGAAGGCGAGT 300
Qy 351 TGAACAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCG 410
Db 301 TGAACAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCG 360
Qy 411 CCTCATCGAGGACGATGATGATGAGCTTGTGTCACCGCAACCAAGTCCCGCTGGGTGC 470
Db 361 CCTCATCGAGGACGATGATGATGAGCTTGTGTCACCGCAACCAAGTCCCGCTGGGTGC 420
Qy 471 AAACGCAATCTTGGTGTTCATGAGTGTGCAAAAGGCTGCTGCTGATTTCCGACGAGCCT 530
Db 421 AAACGCAATCTTGGTGTTCATGAGTGTGCAAAAGGCTGCTGCTGATTTCCGACGAGCCT 480

531 CCACATGTTCCGGCTACATCGGTGGAGCAAAACGACACGTTCTTCCAGTTCCAAATGATGAA 590
 Db |
 481 CCACATGTTCCGGCTACATCGGTGGAGCAAAACGACACGTTCTTCCAGTTCCAAATGATGAA 540
 Qy |
 591 CATCATCAAGGTTGGGCTCAGCTCAGCTCCGGTGTGAGTTTCAGGAATTCATGATCG 650
 Db |
 541 CATCATCAAGGTTGGGCTCAGCTCAGCTCCGGTGTGAGTTTCAGGAATTCATGATCG 600
 Qy |
 651 TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTTACCACGC 710
 Db |
 601 TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTTACCACGC 660
 Qy |
 711 ACTGAAGTCCGTTCATCAAGAAAGGCTCTGTCACCGGACTTGGCGATGAGGGGGCTT 770
 Db |
 661 ACTGAAGTCCGTTCATCAAGAAAGGCTCTGTCACCGGACTTGGCGATGAGGGGGCTT 720
 Qy |
 771 CGCTCTCTCCGTCGGCTCCACCGTTCAGGCTCTTGACCTTATGTTGAGGCAATCGAGAA 830
 Db |
 721 CGCTCTCTCCGTCGGCTCCACCGTTCAGGCTCTTGACCTTATGTTGAGGCAATCGAGAA 780
 Qy |
 831 GGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGTCTTGGAGCTTGTCTTCTCTGAGTT 890
 Db |
 781 GGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGTCTTGGAGCTTGTCTTCTCTGAGTT 840
 Qy |
 891 CTTCAAGGACGGCACTTACCACTTGAAGTGGCCAGCACTCGCAGCTGAGATGCGCAA 950
 Db |
 841 CTTCAAGGACGGCACTTACCACTTGAAGTGGCCAGCACTCGCAGCTGAGATGCGCAA 900
 Qy |
 951 CGTTTACGCTGAGCTGTTGACGGTACCCATCGTCTCTCATCGAGGACCCACTCGAGGA 1010
 Db |
 901 CGTTTACGCTGAGCTGTTGACGGTACCCATCGTCTCTCATCGAGGACCCACTCGAGGA 960
 Qy |
 1011 AGATGACTGGGAGGGTTACACCACTCAGCGCAACATCGGCGCAAGGTTTCAGATCGT 1070
 Db |
 961 AGATGACTGGGAGGGTTACACCACTCAGCGCAACATCGGCGCAAGGTTTCAGATCGT 1020
 Qy |
 1071 TGGCGACGACTTCTTGTTCACCAACCTCGAGCGCTGAGGAGGCGATCGCTTAAGAGGC 1130
 Db |
 1021 TGGCGACGACTTCTTGTTCACCAACCTCGAGCGCTGAGGAGGCGATCGCTTAAGAGGC 1080
 Qy |
 1131 TGGCAACTCCATCTCTGTTAAGGTGAAACAGATCGGTACCTTACCGAGACTTCGAGCG 1190
 Db |
 1081 TGGCAACTCCATCTCTGTTAAGGTGAAACAGATCGGTACCTTACCGAGACTTCGAGCG 1140
 Qy |
 1191 TGTGACATGGCTCACCGGCGAGGCTACACCTCCATGATGTCCACCGTTCCGGTGAGAC 1250
 Db |
 1141 TGTGACATGGCTCACCGGCGAGGCTACACCTCCATGATGTCCACCGTTCCGGTGAGAC 1200
 Qy |
 1251 CGAGGACACCACTTGTGACTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
 Db |
 1201 CGAGGACACCACTTGTGACTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
 Qy |
 1311 TGCTCCAGCACTTCGACCGGTGTCGAAAGTACAAACAGGTTTCTCCGATCGAGCAGCT 1370
 Db |
 1261 TGCTCCAGCACTTCGACCGGTGTCGAAAGTACAAACAGGTTTCTCCGATCGAGCAGCT 1320
 Qy |
 1371 GCTTGGCGAGCGCGGCTTACGAGGTCGAGCGGATTCCTCCACGCTTTCAGGGCTAAT 1430
 Db |
 1321 GCTTGGCGAGCGCGGCTTACGAGGTCGAGCGGATTCCTCCACGCTTTCAGGGCTAAT 1380
 Qy |
 1431 AAAAGCGCTTTTCGACGCCCGGTAA 1455
 Db |
 1381 AAAAGCGCTTTTCGACGCCCGGTAA 1405

RESULT 7
 US-10-781-014-71
 ; Sequence 71 Application: US/10781014
 ; Publication No. US20040180408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompeius, Markus
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Schroder, Hartwig

APPLICANT: Zelder, Oskar
 APPLICANT: Haberhauser, Gregor
 TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
 TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
 TITLE OF INVENTION: PRODUCTION
 FILE REFERENCE: BGI-126PCPN
 CURRENT APPLICATION NUMBER: US/10/781,014
 CURRENT FILING DATE: 2004-02-17
 PRIOR APPLICATION NUMBER: US 09/602,740
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: 60/141,031
 PRIOR FILING DATE: 1999-06-25
 PRIOR APPLICATION NUMBER: 60/143,208
 PRIOR FILING DATE: 1999-07-09
 PRIOR APPLICATION NUMBER: 60/151,572
 PRIOR FILING DATE: 1999-08-31
 PRIOR APPLICATION NUMBER: DE 19931412.8
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19931413.6
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19931419.5
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19931420.9
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19931424.1
 PRIOR FILING DATE: 1999-07-08
 PRIOR APPLICATION NUMBER: DE 19931428.4
 PRIOR FILING DATE: 1999-07-08
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 784
 SEQ ID NOS: 781-1398
 LENGTH: 1398
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (101)..(1375)
 OTHER INFORMATION: RXA00235
 US-10-781-014-71

Query Match 88.5%; Score 1396.4; DB 19; Length 1398;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 51 CGAAACAGATTCGTGCAACAAATTCGGCTTAGAGCTGATTGAACACATTCATCAGTGA 110
 Db |
 1 CGAAACAGATTCGTGCAACAAATTCGGCTTAGAGCTGATTGAACACATTCATCAGTGA 60
 Qy 111 ATAAATTCAGTTCCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAAGT 170
 Db |
 61 ATAAATTCAGTTCCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCAAGT 120
 Qy 171 ATTCGCTCGGAAATTCCTGACTCCCGGCTTAAACCAACCGTTCGAGGACAGAGTTTCT 230
 Db |
 121 ATTCGCTCGGAAATTCCTGACTCCCGGCTTAAACCAACCGTTCGAGGACAGAGTTTCT 180
 Qy 231 GGATGACGTTTCCACGCTGTCGAGGTGTCATTCGCGGCGATCCACCGGCTTCCACGA 290
 Db |
 181 GGATGACGTTTCCACGCTGTCGAGGTGTCATTCGCGGCGATCCACCGGCTTCCACGA 240
 Qy 291 GGCTCATGAGCTCGGTGACGCTGCGATCCCTACCTGGGCAAGGGCTTTTGAAGGCAAGT 350
 Db |
 241 GGCTCATGAGCTCGGTGACGCTGCGATCCCTACCTGGGCAAGGGCTTTTGAAGGCAAGT 300
 Qy 351 TGAACAGCTCAACGAAGAAATTCGCGGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCG 410
 Db |
 301 TGAACAGCTCAACGAAGAAATTCGCGGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCG 360
 Qy 411 CCTCATCGAAGCAATGATCAAGTTGATGCGCCGCAACCAAGTCCCGCTTGGGTGC 470
 Db |
 361 CCTCATCGAAGCAATGATCAAGTTGATGCGCCGCAACCAAGTCCCGCTTGGGTGC 420
 Qy 471 AAACGCAATCCTTGGGTGTTTCCATGGCTGTTGCAAGGCTGCTGCTGATCCCGGAGGCT 530

```

421 AAACCAATCTTGTGTTCATGGCTGTGCAAGGCTGCTGATTCGAGGCT 480
531 CCCACTGTTCCGCTACATCGGTGGACCAAGGCAACGCTTCTCCAGTTCGAATGAGAA 590
481 CCCACTGTTCCGCTACATCGGTGGACCAAGGCAACGCTTCTCCAGTTCGAATGAGAA 540
591 CATCATCAACCGGTGGCTGACGCTGATCCCGGTGTGACGCTTCCAGGAATTCATGATCG 650
541 CATCATCAACCGGTGGCTGACGCTGATCCCGGTGTGACGCTTCCAGGAATTCATGATCG 600
651 TCCCAATCGGTGGAGAGCTTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACCAAGC 710
601 TCCCAATCGGTGGAGAGCTTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACCAAGC 660
711 ACTGAAGTCCGCTCATCAAGGAAAGGCGCTGTGCAACGCGACTTGGCGATGAGGCGGCTT 770
661 ACTGAAGTCCGCTCATCAAGGAAAGGCGCTGTGCAACGCGACTTGGCGATGAGGCGGCTT 720
771 CGCTCCTTCCGCTCGGCTCACAACCGGTGAGGCTTGTGACCTTATCGTTGAGGCAATCGAGAA 830
721 CGCTCCTTCCGCTCGGCTCACAACCGGTGAGGCTTGTGACCTTATCGTTGAGGCAATCGAGAA 780
831 GGCTGGCTTACCCGCGGAGGAGCATCGCTTGTGCTTGGACGCTTGTGCTTGTGCTTGTGCTT 890
781 GGCTGGCTTACCCGCGGAGGAGCATCGCTTGTGCTTGGACGCTTGTGCTTGTGCTTGTGCTT 840
891 CTTCAAGAGCGGCACTTACCACTTCAAGGTTGGCGGAGCACTCCGAGCTGAGATGGCAAA 950
841 CTTCAAGAGCGGCACTTACCACTTCAAGGTTGGCGGAGCACTCCGAGCTGAGATGGCAAA 900
951 CGTTTACGCTGAGTGTGAGCGGTACCCCAATCGTCTCCATCGAGGACCACTTGCAGGA 1010
901 CGTTTACGCTGAGTGTGAGCGGTACCCCAATCGTCTCCATCGAGGACCACTTGCAGGA 960
1011 AGATGATCGGAGGAGGTATACCAACCTTCAAGGTTGGCGGAGCACTCCGAGCTGAGATGGT 1070
961 AGATGATCGGAGGAGGTATACCAACCTTCAAGGTTGGCGGAGCACTCCGAGCTGAGATGGT 1020
1071 TGGCGAGCACTTCTTGTGCTTCAAGGTTGGCGGAGCACTCCGAGCTGAGATGGT 1130
1021 TGGCGAGCACTTCTTGTGCTTCAAGGTTGGCGGAGCACTCCGAGCTGAGATGGT 1080
1131 TGGCAATCTCATCTGTTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTT 1190
1081 TGGCAATCTCATCTGTTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTT 1140
1191 TGTGCAATGAGCTTCAAGGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAG 1250
1141 TGTGCAATGAGCTTCAAGGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAG 1200
1251 CGAGGACACCACTTGTGCTTCAAGGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAG 1310
1201 CGAGGACACCACTTGTGCTTCAAGGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAG 1260
1311 TGTGCAATGAGCTTCAAGGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAG 1370
1261 TGTGCAATGAGCTTCAAGGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAG 1320
1371 GCTTGGCGAGCGGCGGCTTACGAGGTTGAGGCTTCCGAGGCTTCCGAGGCTTCCGAGGCT 1430
1321 GCTTGGCGAGCGGCGGCTTACGAGGTTGAGGCTTCCGAGGCTTCCGAGGCTTCCGAGGCT 1380
1431 AAAAGCGCTTTTCCAGCG 1448
1381 AAAAGCGCTTTTCCAGCG 1398

```

RESULT 8
 US-09-738-626-1085
 ; Sequence 1085, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SBIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATBISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09-738-626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1085
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1085

Query Match      80.8%; Score 1275; DB 9; Length 1275;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 GTGGCTGAAATCATGCAAGTATTCGCTCGCGAAATTCGACTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAAATCATGCAAGTATTCGCTCGCGAAATTCGACTCCCGCGGTAAACCAACC 60
QY 211 GTCGAGGCAAGAGTTTCTCGATGATGACGTTTCCACGTTGTCGAGGTTCATCCCGC 270
DB 61 GTCGAGGCAAGAGTTTCTCGATGATGACGTTTCCACGTTGTCGAGGTTCATCCCGC 120
QY 271 GCATCCACCGCGGTCCACGAGGCTCATGAGCTCGGTGACGAGGCTACCTGAGGCT 330
DB 121 GCATCCACCGCGGTCCACGAGGCTCATGAGCTCGGTGACGAGGCTACCTGAGGCT 180
QY 331 AAGGGCGTTTGAAGGCAAGTTGAAACGTTCAACGAGGAAATCGGCGACGAGCTCGCTG 390
DB 181 AAGGGCGTTTGAAGGCAAGTTGAAACGTTCAACGAGGAAATCGGCGACGAGCTCGCTG 240
QY 391 CTAGAGGCTGACGATCAGGCGCTCATCGACGAGGAAATGATCAAGTTGATGGCACCGCC 450
DB 241 CTAGAGGCTGACGATCAGGCGCTCATCGACGAGGAAATGATCAAGTTGATGGCACCGCC 300
QY 451 AACAAAGTCCCGCTCGGTGCAACCAATCTTGTGTTTCCATGGCTGTTGCAAGGCT 510
DB 301 AACAAAGTCCCGCTCGGTGCAACCAATCTTGTGTTTCCATGGCTGTTGCAAGGCT 360
QY 511 GCTGCTGATTCGCGAGGCTCCCACTGTTTCCGCTACATCGGTGAGGCAACGACACGTT 570
DB 361 GCTGCTGATTCGCGAGGCTCCCACTGTTTCCGCTACATCGGTGAGGCAACGACACGTT 420
QY 571 CTTCCAGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
DB 421 CTTCCAGTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 631 GTTCAGGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
DB 481 GTTCAGGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 691 GCGCGGAGGCTTCAACGAGCTGAGTTCGTTTCAAGGAAAGGCGCTGTCACCGGA 750
DB 541 GCGCGGAGGCTTCAACGAGCTGAGTTCGTTTCAAGGAAAGGCGCTGTCACCGGA 600
QY 751 CTTGGCGATGAGGCGGCTTCCGCTTCCGCTCGGCTCCACCGGTGAGGCTTTCGACCTT 810

```

Db 601 CTTGGGATGAGGGGCTTCGCTCTCTTCGCTCGGCTCAACCGTGAGGCTCTTGACCTT 660
 Qy 811 ATCGTTGAGCAATCGAGAAGGCTGGCTTCAACCCGAGGCAAGACATCGCTCTTGCTCTG 870
 Db 661 ATCGTTGAGCAATCGAGAAGGCTGGCTTCAACCCGAGGCAAGACATCGCTCTTGCTCTG 720
 Qy 871 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCGACCTTACACCTTCAAGGTTGCGGACAC 930
 Db 721 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCGACCTTACACCTTCAAGGTTGCGGACAC 780
 Qy 931 TCGGAGCTGAGATGCAAGCTTTTACGCTGAGCTGTTGACGCTTACCAATCGTCTCC 990
 Db 781 TCGGAGCTGAGATGCAAGCTTTTACGCTGAGCTGTTGACGCTTACCAATCGTCTCC 840
 Qy 991 ATCGAGGACCCACTCGAGGAAGATGCTGGAGGGTTACACCAACCTCACCGCAACCATC 1050
 Db 841 ATCGAGGACCCACTCGAGGAGATGCTGGAGGGTTACACCAACCTCACCGCAACCATC 900
 Qy 1051 GCGCAAGGTTGAGATCGTTGGCGACGACTTCTCGTCAACCAACCTGAGCGCTGAAG 1110
 Db 901 GCGCAAGGTTGAGATCGTTGGCGACGACTTCTCGTCAACCAACCTGAGCGCTGAAG 960
 Qy 1111 GAGGCGATCGTAAGAGGCTGCAACTCGATCTCGTTAAGTGAACAGATCGGTACC 1170
 Db 961 GAGGCGATCGTAAGAGGCTGCAACTCGATCTCGTTAAGTGAACAGATCGGTACC 1020
 Qy 1171 CTCACCGACCTTCGACGCTGCGACATGGCTCACCGCGCAGGCTACACCTTCCATGATG 1230
 Db 1021 CTCACCGACCTTCGACGCTGCGACATGGCTCACCGCGCAGGCTACACCTTCCATGATG 1080
 Qy 1231 TCCACCGCTTCGCGTGAGACCGAGGACACCACTTGTGACCTCGCAGTTGCACTCAAC 1290
 Db 1081 TCCACCGCTTCGCGTGAGACCGAGGACACCACTTGTGACCTCGCAGTTGCACTCAAC 1140
 Qy 1291 TGTGGCAGATCAAGACTGGTGTCAGCAAGTTCGCAAGTTCGCAAGTTCGCAAGTTCGCAAG 1350
 Db 1141 TGTGGCAGATCAAGACTGGTGTCAGCAAGTTCGCAAGTTCGCAAGTTCGCAAGTTCGCAAG 1200
 Qy 1351 CTTCTCCGATCGAGCAGCTGCTTGGCGACGCGGCTTACGCGAGGTCGAGCGGCTTC 1410
 Db 1201 CTTCTCCGATCGAGCAGCTGCTTGGCGACGCGGCTTACGCGAGGTCGAGCGGCTTC 1260
 Qy 1411 CCACGCTTTCAGGGC 1425
 Db 1261 CCACGCTTTCAGGGC 1275

RESULT 9

US-10-728-122A-17724
 ; Sequence 17724, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangeu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Cart, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 17724
 ; LENGTH: 1275
 ; TYPE: DNA
 ; ORGANISM: *Corynebacterium diptheriae*
 US-10-282-122A-17724

Query Match 57.6%; Score 908.6; DB 17; Length 1275;
 Best Local Similarity 82.0%; Pred. No. 4.7e-269;
 Matches 1046; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
 Qy 151 GTGCTGAAATCATGCAGTATTCGCTCGGAAATTCGACTCCCGGGTAACCAACC 210
 Db 1 GTGCTGACATTATGCAGTATTTGCTGCTGAAATCTTGGACTCCCGGGTAACCAACC 60
 Qy 211 GTCAGGCGAGAGTTTCTCGATGACGGTTCCACGSGTGTGCGAGTGTTCATCCGGC 270
 Db 61 GTTGAAGCCGAAGTTTCTTGGATGACGGATCCACGCGCTTGGCGGCTTCTTCCGGT 120
 Qy 271 GCATCCACCGCGCTCCAGGCTCATGAGCTGCGTGACGGTGTGCGAGTGTTCATCCGGC 330
 Db 121 GGTCCACCGGTTTTCAGAGGCTTCCAGAGTTCGCGAGGTTGGCGAGGCTTACCTCGC 180
 Qy 331 AAGGGCGTTTGAAGGCGAGTTGAAACGTCACAAAGAAATCGGCGAGAGTCTCGTGGC 390
 Db 181 AAGGGCGTTCTCAATGCAAGTTTAAACAGTGAACGAGGAAATCGCTGACGCAATCGTGT 240
 Qy 391 CTAGAGCTGAGGATCAGCGCTCATCGACGAAAGATGATCAAGTGTGATGCGACCGC 450
 Db 241 GCAGAAGCCGAGATCAGCGCTTGTATCGATCAGGCGATGATTGCTCTCGAGCGCACTGAG 300
 Qy 451 AACAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGCTGCTTTCGCAAGGCT 510
 Db 301 AACAGTCTCGTCTCGGCGCAACGCTATCTTGGGTGTATCCATCGCGCTAGTAAAGCT 360
 Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCGCTTACATCGGTGCAACCAAGCGCACACCT 570
 Db 361 GCGCAGAGTCTGTGCGCTTACCTTTGTACGCTTACATCGGCGGCTTAAAGCTCACCT 420
 Qy 571 CTTCCAGTTTCAATGATGAACATCATCAACGTTGGCGCTCACGTTGATTCGGTGTGAC 630
 Db 421 CTTCCAGTTTCTATGATGAACATTTGTTAAACGTTGGCGCACACGCTGACTCCGGGCTTAT 480
 Qy 631 GTTCAGGAATTCATGATCGCTCCATCGGTGCAAGACCTTCTCTGAGGCTCTCGGCAAC 690
 Db 481 GTTCAGGAGTTTCAATGATGCTTATCGGTGCGAGTCTTCTCTGAGGCTCTCGGCAATG 540
 Qy 691 GCGCGAGGCTTACCAACGCACTCAAGTTCGCTATCAAGGAAAGGCGCTCTGCAACCGA 750
 Db 541 GGTGCAAGGTTCTACCACTCTTGAAGTTCGTTAAGTCAAGGAGACTTTTCCACCGGC 600
 Qy 751 CTTGGCGATGAGGCGGCTTTCGCTTCTTCCGTGCGCTCCACCGCTGAGGCTTTCGACCT 810
 Db 601 CTCGGCGCAAGAGTGGTTTCGACCTTCTGTTGAGTCCACCAAGGCGAGCTCTCGACCTC 660

Db 4372589 CGAGGCGCTGGCCCGGGGATCGAGAGGGCTCCGCCAAGCCCTGCTCGTCAAGGTCA 4372530
QY 1157 ACCAGATCGGTACCTTCAACGAGACCTTCGACGCTGTGACATGGCTCAACCGCGAGGCT 1216
Db 4372529 ACCAGATCGGTTCGCTGACCGAGACCTTGAGCGCGTTCGAGCTGGCCGAGCGAAGGCT 4372470
QY 1217 ACACCTCATGATGTCCTCAACGCTTCGCGTGTGAGACGAGGACACCAATTCGTGACCTCG 1276
Db 4372469 TCAAGTGCATGATGTCCCAACGCTCCGCGGAGACCGAGGACGTCACCATCGCGGACCTCG 4372410
QY 1277 CAGTTGCACTCAACTGTGTGCGAGATCAAGACTGTGTCTCCAGCAGGTTCCGACCGGTGCG 1336
Db 4372409 CGTGCCTGGAATGCGGTTCAGATCAAGACGCGGCCCTCGACCGGTGCG 4372350
QY 1337 CAAAGTACACAGCTTTCGCGATCGAGCAGTGTCTGGCGACGCGCGCTGTAGCGAG 1396
Db 4372349 CCAAGTACAAACAGCTGTCTGCGATCGAGGAGATCTTCGACGACGCGCGGAGTAGCGG 4372290
QY 1397 GTGCGAGCGCATTCACAGCGTTTC 1420
Db 4372289 GCGCTCGGCGTTCGCCCGGTTTC 4372266

RESULT 11

US-10-156-761-3520
; Sequence 3520, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3520
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1284)
US-10-156-761-3520

Query Match 44.2%; Score 697; DB 15; Length 1284;
Best Local Similarity 72.3%; Pred. No. 8.3e-204;
Matches 920; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

QY 151 GTGGGTGAATCATGACGATATTCGCTCGCGAAATTCGACTCCCGGGTACCCCAACC 210
Db 1 GTGCGGTTCATCGACTGCTGCTAGCCCGGAAATCTTGACTCCCGAGGCAACCCACG 60

QY 211 GTGAGGAGAGGTTTCTCTGATGACGCTTCCACGCTGCGAGGTGTTCCATCCGCG 270
Db 61 GTGAGGTGAGGTTCGGCTTCACGACGCGACGACGAGGTGTCGCGCGCTCCGTCGCGC 120

QY 271 GCATCCACCGCGGTCCACGAGGCTCATGACTGCGGTGACGCGTG---GGCATCGCTACCTG 327
Db 121 GCCTCCACCGGTGCTTCGAGGCGCATCGAGCTCCGCGACGGTGACCCCAACCGTTACCGAG 180

QY 328 GGCAGAGGCGTTTGAAGGCAAGTTGAAACGTCACGAGAAATCGGCGACGAGTCGCT 387
Db 181 GGCAGAGGCTGCGAAGAGCCGCTCGCGCTCATCGAGCAGATCGGCGCGGAGCTCGTC 240

QY 388 GGCCTAGAGGTGACGATCAGCGCTCATCAACGAGCAATGATCAAGCTTGTATGGCAC 447
Db 241 GGTACGAGCGCACCGAGCGGCTGATCGACAGGCGATGTTGACGCTGAGCCAC 300
QY 448 GCCACAGTCCCGCTGGGTGCAACGCAATCTTGTGTTTCCATGGCTGTTGCAAG 507
Db 301 GACAAACAGGCTCGCTCGGCGCAACGCAATCTTGTGCGCTCTCCCTCGCGCTCGCGAC 360
QY 508 GCTGTGTGATTCGCGAGGCTCCCACTGTTTCCGCTACATCGGTGAGCAAAACGACAC 567
Db 361 GCGCTCCGAGGCGACGACCTCCGCTCTTCCGCTACCTGGGCGGCGGACGCGAC 420
QY 568 GTTCTTCAGTTCAAATGATGAACATCATCAACGCTGGCGCTCACGCTGATCTCGGTT 627
Db 421 CTGTGCGCTTCGCTGATGAACATCTCGTGAACGCGCTCCGACGCGGACTCCAAACGT 480
QY 628 GACGTCAGGAAATCATGATCGCTCCAAATCGGTGCGAGACCTTCTCTGAGGCTCTCGC 687
Db 481 GACATCCAGGATTCATGATCGCCCGGATCGGCGGAGTCTTCTCCGAGGCGCTCGGC 540
QY 688 AACGCGCGAGGCTTACACGCACTGAAGTCCGTCATCAAGGAAAGGGCTGTCCAC 747
Db 541 TGGGCGCGAGGCTTACCAACACCTCAAGAGGTGCTGAAGACCAAGGGCTGTCCAC 600
QY 748 GGAATTGCGATGAGGCGGCTTCGCTCTTCTCCGCTCCACCCGTGAGGCTCTTGAC 807
Db 601 GGCCTCGGCGAGGCGGCTTCGCCCGAACTCGGAGTCAACCGCGCGCTCGAC 660
QY 808 CTTATCGTTGAGGCAATCGAAGGCTGGCTTCAACGAGCGGCTTCAACCGAGCAAGGATCGCTT 867
Db 661 CTCATCATCGAGGCGCATCAAGCGGCGGTTTACATCCCGGCGGAGCAGATCGCGCTCGC 720
QY 868 CTGAGGCTGCTTCTCTGAGTTCTTCAAGGAGCGGCTTCAACCGAGGCTTCAACCGAGTGGCCAG 927
Db 721 CTCAGCTCGCGCTCGAGTTCTTCAAGGAGCGGCTTCAACCGAGGCTTCAACCGAGTGGCCAG 780
QY 928 CACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGAGCGCTTACCAATCGCTC 987
Db 781 GGTCCGCGCGGAGTACCGGAGTCTTACGAGGAGTCTTCCGCGTACCGGCTCGTC 840
QY 988 TCCATGAGGAGCCACTGCGAGGAGATGACTGGGAGGTTTACCAACCTCACCGCAAC 1047
Db 841 TCCATGAGGAGCCGCTGTACGAGGAGGACTGGGCGGCTTGAAGGTTTATCACCGGACAA 900
QY 1048 ATGCGCGACAGGTTTCAAGTGGGCGAGGCTTCTTCTGTCACCAACCTTACCGGCGCTG 1107
Db 901 CTGGGCGACAGGTTCCAGATCGTGGGCGAGGCTTCTTCTGTCACCAACCGGAGGCGCTG 960
QY 1108 AAGGAGGCGCATCGCTAAGAGGCTGCGCAACTTCCATCTCTGTTAAGGTGAACAGATCGGT 1167
Db 961 GCGCGCGCATCGAGGAGGCTCGGCCAAGCGCTTCTGCTGCTCAAGTCAACAGATCGGT 1020
QY 1168 ACCCTCACGAGACTTTCGAGCTGTGCAATGCTTCAACGCTCACCGGAGGTTTACCTTCAATG 1227
Db 1021 TCGCTGACCGAGACCTTGGAGCTGAGCTGGCGGCGGAGGCTTCAAGTTCATG 1080
QY 1228 ATGTCCACCGTTCGCTGAGGAGGAGGACACCACTTCTGCTGAGCTGCGGAGTTGCACT 1287
Db 1081 ATGTCCACCGCTTCGCGAGGAGGAGGAGGAGTCAACATCGCGGAGCTTCTGCGCTGCGCGT 1140
QY 1288 AACTGTGGCGAGTCAAGACTGCTGCTCCAGCAGCTTCCGAGCGGCTGCGCAAGGTACAAAC 1347
Db 1141 AACTGCGGTGATCAAGACGCGGCGCGCGCTTGGACCGCTTGGCGAGTACAAAC 1200
QY 1348 CAGCTTCTCGCATCGAGCAGCTTCTTGGCGAGCGCGCGCTTACGAGGTTGCGAGCGCA 1407
Db 1201 CAGCTGCTGGCATCGAGGAGATCTTCGACGAGCGCGCGGAGTACGCGCGCTCGCGC 1260
QY 1408 TTCCACGCTTTC 1420
Db 1261 TTCCCGGCTTTC 1273

RESULT 12

US-10-369-493-32027
; Sequence 32027, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 32027
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Thermobifida fusca
US-10-369-493-32027

Query Match 43.1%; Score 679.8; DB 17; Length 1269;
Best Local Similarity 71.0%; Pred. No. 1.7e-198;
Matches 900; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY	152	TGGCTGAAATCATGCAAGCTATTTCGCTCGGAAATTCGACCTCCCGCGGTAAACCCACCG	211
DB	2	TGGCGTCGATCAGGCGATGACACGACGGGAGATTCTCGACTCCCGCGGTAAACCCGACGG	61
QY	212	TCGAGCGAGAGGTTTTCTCTGGATGACGGTTCCACAGGTGTCCAGAGTGTTCATCCGCGC	271
DB	62	TCGAGTTCGAGTTCGCACTCGACGACGGACGATTTGCCCGGGGCTGTCCCACTGGCG	121
QY	272	CATCACCGGCTCCACGAGGCTCATGAGCTCGGTGAGCGTGGCGATCGCTACCTGGGCA	331
DB	122	CCTCACCGGCTCAGTTTCGAAGCGGTTCGAATTCGCGGATGTGGCGACGGTTACCGCGTA	181
QY	332	AGGGCGTTTTGAAGCGAGTTGAAGACGTCACAGAGAAATCGGACGAGCTCGCTGGCC	391
DB	182	AGGGGTGGAGAGGCGCTCGCGCGGTCCACAGAGAAATTCGCAACAGATCTGGGT	241
QY	392	TAGAGGCTGACGATCAGCGCTTCATCGACGAGCAATGATCAAGCTTGATGCAACCGCA	451
DB	242	ACGAGGCGAGGACGAGCGGCTCGTCCGACAGCGCTCATCACCTGGACGTTACCCCG	301
QY	452	ACAAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGCTGTTCGAAAGGCTG	511
DB	302	ACAAATCGCGCTGGGCGCCAAACGCAATCCTGGGTGTCTCCGTGGCGGTGGCCAGGCG	361
QY	512	CTGCTGATTTCGCGAGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACAGCTTC	571
DB	362	CGCGGAGAGCGGACCTCGCGCTCTTCGGTACTCTGGCGGCGGCAAGCGCCACGTGC	421
QY	572	TTCCAGTTCCATGATGACATCATCAAGTGGGCTCAGCTGACCTCCGGTGTGACG	631
DB	422	TGCGGTGCGATGATGAATTCCTCAACGGGGTGGCGACGCGGACCAACGATCGACA	481
QY	632	TTCAAGATTCATGATCGCTCCATCGGTGCAAGACCTTCTGAGGCTTCGCGAACG	691
DB	482	TCCAGAGTTTCATGATCGCCCCCATCGGGGCGGAGAGCTTCGGGAGCGCTGGCGTGG	541
QY	692	GGCGGAGGTTTACCACGACCTGAAAGTCCGTTCATCAAGGAAAGGGCTGTCCACCGGAC	751
DB	542	GGCGGAGGTTTACCACCTGCTCAAGGCGGTGCTCAAGCCCAACGCTTGCGCCACGGTG	601
QY	752	TGGCGATGAGGGCGGCTTCGCTCTTCGTCGGCTCCACCGGTGAGGCTCTTGACCTTA	811
DB	602	TCGCTGACGAGGGCGGGTTCGCGCCCAACCTGGACAGCAACCGCGCGCTCGACCTGA	661
QY	812	TGCTTGGGCAATCGAGAAGGCTGGCTTCACCCCGGACGAGACATCGCTCTTGTCTGG	871

DB	662	TCAGCGAAGCCATCACGAAGCGGGCTTCACTGGGCCAGGACATCGCCCTGGCTGG	721
QY	872	ACGTTGCTTCTCTGAGTCTTCAAGGACGGCACTTACCACTTCGAAGGTGGCCAGCACT	931
DB	722	ACGTGGCCGCCACCGAGTTCTACCGGACGGCGCTTACCACTTCGAAGGTGGCCAGCA	781
QY	932	CCGACGCTGAGATGGCAAAAGTTTACGCTGAGCTCGTTGACCGGTACCCAATCTCTCCA	991
DB	782	CGCGGAGGAGATGGCGCCTACTACACCGAGCTGTGCTGAGTCTTACCGCTGTGTGGA	841
QY	992	TCGAGGACCCACTCCAGGAGAGATGCTGGGAGGTTTACCAAACTTCAACCGAACCATCG	1051
DB	842	TCGAGGACCCGCTCAGCGAGGAGACTGGGCGGGTGGAAAGGGCTCACCGAGTCCCTCG	901
QY	1052	GCACAAAGGTTTCAGATCGTTGGGAGGAGCTTCTTCGTCACCAACCTTGGAGGCTGAAG	1111
DB	902	CGGACCGGCTGCAACTGGTCGGGAGACCTGTTTCGTCACCAACCTTGGAGGCTGAGC	961
QY	1112	AGGCGATCGCTAAGAGGCTGCCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCC	1171
DB	962	CGGTATCGAGGAGGCGCTGCCAACTCTGCTGTGCTCAAGGTGAACAGATCGGTACGC	1021
QY	1172	TCACCGAGACTTCGACGCTGTGACATGCTCACCGCGAGGCTACACCTCCATGATGT	1231
DB	1022	TCAGCGAGACTTTGGACGCGGTGTGCTGGCGGAGCGCAACCGGTACACCGCATGATCA	1081
QY	1232	CCACCGGTTCCGCTGAGACGCGAGGACACCACTTGTGTCGACCTCGGAGTTCGACTCACT	1291
DB	1082	GCACCGGTTCTGCTGAGACCGGAGACCACTGCTGCGGACATCGCGGTGGCGGCAACT	1141
QY	1292	GTGCCAGATCAAGACTGTGCTCCAGCACTTCCGACCGCTGCGAAAGTACAAACCGC	1351
DB	1142	CGGGGAGATCAAGCGGTTGCCCGCGCGGAGCGGCTGCGCAAGTACAAACCGC	1201
QY	1352	TTCTCGCATCGAGAGCTGCTTGGGAGCGCGGCTGTACGAGGTCGAGGCGCATTC	1411
DB	1202	TGCTGCGATCGAAGAGAGCTCGAGACGCGCGCTCTACGCGGAGCAACGCGTTC	1261
QY	1412	CACGCTT 1418	
DB	1262	CGGTTT 1268	

RESULT 13

US-10-282-122A-25949
; Sequence 25949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25949
; LENGTH: 1287
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25949

Query Match
  41.0%; Score 647; DB 17; Length 1287;
Best Local Similarity 69.9%; Pred. No. 2.2e-188;
Matches 888; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

QY 151 GTGGCTGAATCATGCACGTATTTCGCTCGCGAAATTCGACCTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAATCATGCACGTATTTCGCTCGCGAAATTCGACCTCCCGCGGTAAACCGACA 60

QY 211 GTGAGGCGAGAGGTTTCTCGATGACGTTTCCACGCTGTCGAGGTGTTCCATCCCGGC 270
DB 61 GTGAGGTCGAGATCGCCTCTGACCGAAGGATGTTCCGCGCGGCGGTGCGTCCGGG 120

QY 271 GCATCAACCGGCTTCCAGAGGCTCATGAGCTGCGTACGCTGCGGATCGCTACTCTGGGC 330
DB 121 GCGTCAACCGGCTTCCAGAGGCTCATGAGCTGCGTACGCTGCGGATCGCTACTCTGGGC 180

QY 331 AAGGCGTTTTCAGAGCAGTTTGAAGCTTCAACGATCAACGAGAAATCGGACGAGCTCGCTGC 390
DB 181 AAGGCGTTCAGAGCAGTTTGAAGCTTCAACGATCAACGAGAAATCGGACGAGATCGGCGCGGCTGATCGGG 240

QY 391 CTAGAGGCTGACGATCAGCGCTCTACGACGAGCAATGATCAAGCTTGTAGGCAACCGCC 450
DB 241 CTGACGCGGACGACGAGCGCTTGTGGACGAGGCTGCTGATGAGCGGCAACCGCC 300

QY 451 AACAGTCCCGCTCGGTGCAACGCAATCTCTGTTGTTTCCATGCTGTTTCCAAAGGCT 510
DB 301 GACAAATCGAGGCTGCGCGGCAACGCAATCTCTGTTGTTTCCATGCTGTTTCCAAAGGCT 360

QY 511 GCTGCTGATTCGCGAGGCTTCCACTGTTCCGCTATACGCTGAGGACCAACGCAACGCT 570
DB 361 GCGCGGATTCGCGAGGCTTCCACTGTTCCGCTATACGCTGAGGACCAACGCAACGCT 420

QY 571 CTTCAGTTCCTCAATGATGAACTATCAACGCTGCGCTCAGCTGAGCTTCCGCTGTTGAC 630
DB 421 CTGCGGCTCCGATGATGAACTATCAACGCTGCGCTCAGCTGAGGACCAACGCTGAGC 480

QY 631 GTTCAGGAATTCATGATCGCTTCAATCGGTGAGAGACCTTCTCTGAGGCTTCCGCAAC 690
DB 481 ATCCAGGAGTTTATGTTGCGCGCGATCGCGCGCGGCTTTCGCGAGGATTCGCTGTTG 540

QY 691 GCGCGGAGGTTTATCAACGCTATGAGTCCGCTATCAAGGAAAGGCTGTTCCACCGGA 750
DB 541 GGTGCGGAGGTTTATCACTCGTGAAGTCTGCTGAGAGGAGGCTGAGCAACCGCGC 600

QY 751 CTTCGCGATGAGGCGGCTTCCCTTCTTCGCTGCGCTCCACCGGAGGCTTTCACCTT 810
DB 601 CTGCGGAGGAGGCGGCTTTCGCTTCCGCTGAGGCTGAGGCTTTCGCTTTCGCTT 660

QY 811 ATCGTTGAGGCAATCGAAGAGGCTGCTTCAACCGGAGGAGGAGGAGGAGGAGGAGG 870
DB 661 ATCGCGCGGCGCATCGAATCGCGCGGCTTCAACCTCGGAGGAGGAGGAGGAGGAGG 720

871 GACGTTGCTCTCTGAGTTCTTCAAGGACGCGAC ---CTACCATTCTTGAAGGTTGGCCAG 927
721 GACGCGGCGGCGCACCGAGTTCTACAGCGACCGGCTACAGTTTGAAGGCGAGCACC 780
928 CACTCCGAGCTGAGATCGCAAAACGTTTACGCTGAGCTCGTTGACGCGTACCCAAATCGTC 987
781 CGCACCGCGGAGCAGATGCGCGAGTTCTACGCGCGGCTGCTCGCGCGGTATCCGTTGGTG 840
988 TCCATCGAGGACCCACTGACGAGGAGATCACTGGGAGGTTTACCAACCTCACCGCAACC 1047
841 TCCATCGAAGACCCGCTGTCCGAAGATGATTGGGACGCTGGGCGGCGCTGACCGCGTCG 900
1048 ATCGCGCAAAAGTTTCAGATCGTTGGCGACGACTTCTTTCGTCAACCAACCTGAGCGCTG 1107
901 ATCGCGGACCCGCTGACGCTGTCGCGGACGACGCTTCTTCTGTCACGAAACCCCGAAGCCTG 960
1108 AAGGAGGCGATCGCTTAAGAGGCTGCCAACTCCATCTCTGTTTAAAGGTGAACCCAGATCGGT 1167
961 GAAGAGGCTATCGAAGAGGCGTTCGAAATGCGTTGCTGTTCAAGGTGAATCAGATCGGC 1020
1168 ACCCTCAGGAGACTTCGACGCTGTCGACATGCTACCGGCGAGGCTACACCTCCATG 1227
1021 ACCTGACGAGACGCTGACGCTGCTGCGCTGCGCTGCGGCGGCTACCGCGGCTACCGCGATG 1080
1228 ATGTCCCAACCGTTCCGCTGAGACGAGGACACCACTTGTGCTGACCTCGCAGTTGCACATC 1287
1081 ATGAGCCACCGGAGGCTGAAACCGAGGACACCACTGCGCGACCTGCGCGTGGCGGCTC 1140
1288 AACTGTGCGCAGATCAAGACTGCTGCCAGCAGCTTCCGACCGCTGTCGCAAGGTACAAAC 1347
1141 GCGAGCGGCGAGATCAAGACCGGAGCGCGCGCGCGAGCGGCTGCGCAAGGTACAAAC 1200
1348 CAGCTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGGCTTACGAGGTCGAGCGCA 1407
1201 CAGCTGCTCGGATCGAGGAGGCGCTGCGCGACGCGCGCTTACGCGGCGGACCTGGCC 1260
1408 TTCCCAAGCTT 1418
1261 TTCCCGCGGTT 1271

RESULT 14
US-10-080-170-649
; Sequence 649, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 649
; LENGTH: 45191
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-649

Query Match
  40.1%; Score 632.6; DB 15; Length 45191;
Best Local Similarity 68.4%; Pred. No. 2.5e-183;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

QY 140 AGGAGGCGCAGTGGCTGAAATCATGACGTTATTCGTCGCGAAATTCCTGACTCCCGCG 199
DB 19931 AGGAGAACCCAGTCCGCTTATCGAGCAGGTTAGGCGCGGAGAGATCTCTGATTTCCCGCG 19990
200 GTAACCAACCGTCGAGCAGAGGTTTCTGATGAGGTTTCCACGCGGTTCGAGGTG 259
```

Db 19991 GCACCCGACGCTGGAGTTCGAGTGGCGCTTATCGACGGGACATTCGCCCGGCGCGG 20050
Qy 260 TTCCATCCGCGCATCCACCGCGTCCACGAGGCTCATGAGCTCGTGAAGTGGCGATC 319
Db 20051 TGCCGTGCGGCGCTCGACCGGGAGCACAGGCGCTCGAGTTGGCGCAGCGCGCGATC 20110
Qy 320 GCTACTCGGCAAGCGGCTTTTGAAGGCGAGTTGAAACGTCACAAAGAAATCGCGACG 379
Db 20111 GCTACGGCGCAAGCGGTGCAAAAGCGTGCAGGCTGTTCTTGATGAGATCGGCGCG 20170
Qy 380 AGCTCGCTGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
Db 20171 CGGTCAATCGACTCAACGCGAGCACGAGGATGGTGCACAGCGCTGGTGACCTAG 20230
Qy 440 ATGGCACCGCCAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTG 499
Db 20231 ACGGCAACCCCGACAAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 20290
Qy 500 TTGCAAGGCTGCTGCTGATTCGCGAGCGCTCCCACTGTCGCTACATCGTGGTGAACAA 559
Db 20291 TTGCCAAGCGCGCGGATTCGCGGAGCTGCCGTTGTTCCGTTATGTCGGGGGCCAA 20350
Qy 560 AGCACACGTTTCTTCAGTTCCAAATGATGAACATCATCAACGGTGGCGCTCACGCTGACT 619
Db 20351 ACGGCAATCTTGCCTGACGATGATGAACATCTCAACGGCGGCGCACACGCCGATA 20410
Qy 620 CGGTGTTGACGTTTCAAGAAATTCATGATCGCTCCATCGGTGACAGACCTTCTCTAGG 679
Db 20411 CGGTGTCGACATTCAGAGTTTCAATGAGTGGCGGCAATTTGGCGCGCCACGCTTCTCGAGG 20470
Qy 680 CTCTCCGCAACGGCGCGAGGTCTACACGACACTGAGTCCGTATCAAGGAAAGGGCC 739
Db 20471 CGTTGGCTGGGCGCTGAGTGTATACACGGCTCAAGTCCGTCTGAAAGAGGGGC 20530
Qy 740 TGTCCACCGACTTGGCGATGAGGCGGCTTGGCTCTTCCGTCCGTCCACCGTGAAG 799
Db 20531 TGTCCACCGGCTTGGCGACGAAGGCGCTTGGCGCGGATGTGGCGGACACCGCGG 20590
Qy 800 CTCTTGACCTTATCGTTGAGCAATCGAAGAGGTGGCTTCAACCGAGGCAAGACATCG 859
Db 20591 CGTTGGACCTATCAGCGCGGCGCATCGAGTGGCGGCTTGGCACCGCGCGCGACGTG 20650
Qy 860 CTCTTGCTCTGAGTGTCTTCTGAGTCTTCAAGGAGCGCAC---CTACACCTTCG 916
Db 20651 CGCTGGCTTGAACGGCGGCGCACGAGTCTTCAACGACGCGACCGCTACGCTTCG 20710
Qy 917 AAGTGGCGACGACTCCGACGCTGAGATGGCAAGCTTTAGCTGAGCTCGTTGACGCGT 976
Db 20711 AGGGCACACCCGTTACCGACACGATGACCGAGTTCTACGCGGCGCTGCTGGCGCT 20770
Qy 977 ACCCAATCGTCTCCATCGAGGACCCCATGCGAGGAAGATGATGGAGGGGTTACACCAAC 1036
Db 20771 ACCCGCTGGTGTGATCGAAGACCCACTGTTCGAAGACGATTTGGGACGCGCTGGCGCGC 20830
Qy 1037 TCACCGCAACCATCGGCGACAAGGTTGAGTCTGTTGGCGAGCTTCTTGTTCACCAAC 1096
Db 20831 TGAAGGCTCGATCGGTGACCGGCTGCAATCTGCGGAGCAATCTTTGTACCAATC 20890
Qy 1097 CTGAGCGCTGAAGAGGCGATCGCTAAGAGGCTGCCAATCCATCCTGTTAAGGTGA 1156
Db 20891 CCGAGCGGCTCGAGGAGGCGATCGAAGCGGCGGTGGCAATTCGTTGCTGCTCAAGGTGA 20950
Qy 1157 ACCAGATCGGTACCTCACCGAGACCTTCGACGCTGTGACATGGCTCACGGCGAGCT 1216
Db 20951 ACCAGATCGGAGCTTGAACGAGACACTCGACGCGGTTCACGCTGGCTTCAACAGGCGAT 21010
Qy 1217 ACACCTCCATGATGCCACCGTTTCGGTGAAGACCGGAGCACCAACATTTGCTGACCTCG 1276
Db 21011 ACCGACGATGATCAGTCAACCGCAGTGGCGAGACGAGGACACCATGATGCCGACCTCG 21070
Qy 1277 CAGTTGACCTCACTGTGGCGAGATCAAGACTGTGTCTCAAGCAAGTTCGACCGGTGCG 1336
Db 21071 CGGTGGCCATCGGACGCGGCGAGATCAAGACGGGCGCGCTGCTCGCAGTGAAGCGCGT 21130

Qy 1337 CAAAGTACAAACCAAGCTTCTCCGATCGAGCAGCTGTGTGGCAGCGCGCGTCTACGCG 1396
Db 21131 CAAATATACAAACCAAGCTTCTCCGATCGAGCAGCTGTGTGGCAGCGCGCGTCTACGCG 21190
Qy 1397 GTCGACGCGATTCACCAAGCTTTCAGGGCTTAATAAAGCGCT 1439
Db 21191 GCGACCTGGCAATTCCTCCGTTCCGTCGCGAGACGAATAGGT 21233

RESULT 15

US-10-080-170-649
; Sequence 649, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 649
; LENGTH: 45191
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-649

Query Match 40.1%; Score 632.6; DB 19; Length 45191;
Best Local Similarity 68.4%; Pred. No. 2.5e-183;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGCTGAATCATGCACGTATTGCTCGGAAATTCCTCGGAAATTCCTCGACTCCGCG 199
Db 19931 AGGAGAACCCAGTGGCGCATTTATCGAGCAGGTATTAGGCCCCGAGAGATCTCTGATTTCCCGCG 19990
Qy 200 GTAACCCAAACCGTCGAGGCGAGAGGTTTCTCGATGACGCTTCCCAACGCTGTCGACAGGTG 259
Db 19991 GCACCCGAGCGTGGAGGTGAGGTGGCGCTTATCGACGGGACATTCGCCCGCGCGCGCG 20050
Qy 260 TTCCATCCCGCGCATCCACCGCGCTCCACGAGCTCATGAGCTGCGTGAAGTGGCGGATC 319
Db 20051 TGCGTGGCGCGCTCGACCGGGAGCACGAGCGCTGAGTTGCGCGACGCGCGCGATC 20110
Qy 320 GCTACTCGGCGAAGGCGCTTTTGAAGGCGAGTTGAAAGCGAGTTGAAAAAGTCAACGAGAAATCGCGCAG 379
Db 20111 GCTACGGCGCAAGGCGGTGCAAAAAGCGCTGCAAGCTGTTCTTGATGAGATCGGCGCGG 20170
Qy 380 AGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
Db 20171 CGGTCAATCGACTCAACGCGGAGCACGAGGATGGTGCACCGCGCTGGTGACCTAG 20230
Qy 440 ATGGCACCGCCAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTG 499
Db 20231 ACGGCAACCCCGACAAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 20290
Qy 500 TTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAGCCAA 559
Db 20291 TTGCCAAGCGCGCGGATTCGCGGAGCTGCCGTTGTTCCGTTATGTCGGGGGCCAA 20350
Qy 560 ACGCACACGTTTCTTCAGTTTCCAAATGATGAACATCATCAACGGTGGCGCTCACGCTGACT 619
Db 20351 ACGGCAATCTTGCCTGACGATGATGAACATCTCTCAACGGCGGCGCACACGCCGATA 20410
Qy 620 CGGTGTTGACGTTTCAAGAAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGG 679
Db 20411 CGGTGTCGACATTCAGAGTTTCAATGAGTGGCGGCAATTTGGCGCGCCACGCTTCTCGTAGG 20470
Qy 680 CTCTCCGCAACGGCGCGAGGTCTACACGACACTGAGTCCGTATCAAGGAAAGGGCC 739

Db	20471	CGTTGGCTGGGGCGCTGAGGTGTACACCGCGCTCAAGTCGCTCTGAAAAAGGAGGGGC	20530
Qy	740	TGTCCACCGGACTTGGCGGATGAGGGCGCTTGGCTCTTCCTCGGCTCCACCCGTGAGG	799
Db	20531	TGTCCACCGGCTTGGCGGATGAGGGCGCTTGGCTCTTCCTCGGCTCCACCCGTGAGG	20590
Qy	800	CTCTTGACCTTATCGTTGAGGCAATCGAGAAGGCTGGCTTACCCCGAGGCAAGGACATCG	859
Db	20591	CGTTGGACCTGATCAGCCGGGCCATCGAGTCGGCGGGCTTGGACCCCGCGCGGACGTGG	20650
Qy	860	CTCTTGCTCTGAGCGTTGCTTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG	916
Db	20651	CGCTGGCCCTGAGCAGCGCGGCCACCGAGTTCTTACCGACGGCACCGGCTACGCTTTCG	20710
Qy	917	AGGTGGCCAGCACTCCGCGAGCTGAGTGGCAAAAGTTTACGCTGAGCTCGTTGAGCGCT	976
Db	20711	AGGGCACCAACCGGTACCGCAGACCAAGTACCGAGTTCTACCGGGCTTGTCCGGCGCT	20770
Qy	977	ACCCAATCGTCTCCATCGAGGACCCACTGCAGGAAGATGACTGGGAGGGTTACACCAACC	1036
Db	20771	ACCCGCTGGTGTGATCGAGACCCACTGTCCGAGACGATTTGGACGGCTGGGCGCGGC	20830
Qy	1037	TCACCGCAACCAATCGGCGCAAGGTTACAGTCGTTGGCGAGCACTTCTTCGTCAACCAACC	1096
Db	20831	TGACGGCTCGATCGGTGACCGGGTGCAAAATCGTGGCGAGCACATCTTTGTCAACCAATC	20890
Qy	1097	CTGAGGGCTGAAGAGGGCATCGCTAAGAGGCTGCCAACTCCATCTCTGGTTAAGGTGA	1156
Db	20891	CCGAGCGGCTCGAGGAGGGCATCGAAACGGGGCGTGGCAAAATCGTTGTGTCAAGGTGA	20950
Qy	1157	ACCAGATCGGTACCTTCAACGAGACCTTCCGAGACGCTGCGACATGGCTCACCGCGAGGCT	1216
Db	20951	ACCAGATCGGACGTTGACCGGAGACACTCGACCGGTACGCTGGCTTACCAACGGCGAT	21010
Qy	1217	ACACCTCCATGATGTCCACCGTTCCGGTGAGACCGAGGACACCACTTGTCTGACCTCG	1276
Db	21011	ACCGCACGATGATCAGTCACCGCAGTGGCGGAGACGGAGGACACCATGATCGCCGACCTCG	21070
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAGACTGGTGCTCCAGCAGTTCCGACCGGTGCG	1336
Db	21071	CGGTGGCCATCGGACAGCGGGCAGATCAAGACGGGGCGGCTCTGCTGAGTGAGCGGTCG	21130
Qy	1337	CAAGGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGACCGCGGCTTACGCGAG	1396
Db	21131	CAAAATACAAACAGCTGCTGGCGATCGAAGAGGGCTTGGCGACCGCGGCTTACGCGG	21190
Qy	1397	GTCCGAGGGCATTCACCGCTTTCAGGGCTAAATAAAGCGCT	1439
Db	21191	GGGACCTGGCATTTCTCGGTTCCGCTGCGAGACGAAATAGT	21233

THIS PAGE BLANK (user)

THIS PAGE BLANK (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 12:13:58 ; Search time 1026 Seconds
(without alignments)
171.954 Million cell updates/sec

Title: US-10-728-947-2
Perfect score: 2155
Sequence: 1 VAEIMHVFAREILDSRGNPT.....QLLGAGVYAGRSAPFRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No: is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	9 US-09-860-768-2	Sequence 2, Appli
2	2155	100.0	425	16 US-10-728-947-2	Sequence 2, Appli
3	2152	99.9	425	9 US-09-738-626-4585	Sequence 4585, Ap
4	2151	99.8	425	9 US-09-860-768-4	Sequence 4, Appli
5	2151	99.8	425	16 US-10-728-947-4	Sequence 4, Appli
6	2149	99.7	425	16 US-10-781-014-72	Sequence 72, Appli
7	2149	99.7	425	17 US-10-494-836-8	Sequence 8, Appli
8	1917	89.0	425	15 US-10-282-122A-53908	Sequence 53908, A
9	1566	72.7	423	15 US-10-369-493-8340	Sequence 8340, Ap
10	1558.5	72.3	429	15 US-10-282-122A-62133	Sequence 62133, A
11	1555.5	72.2	428	14 US-10-156-761-11070	Sequence 11070, A

ALIGNMENTS

RESULT 1

US-09-860-768-2
; Sequence 2, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mckel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-2

12	1551.5	72.0	429	15	US-10-282-122A-64490	Sequence 64490, A
13	1551.5	72.0	439	15	US-10-282-122A-62696	Sequence 62696, A
14	1501.5	69.7	447	15	US-10-282-122A-63682	Sequence 63682, A
15	1418	65.8	423	15	US-10-369-493-11696	Sequence 11696, A
16	1408.5	65.4	426	15	US-10-369-493-16911	Sequence 16911, A
17	1406	65.2	429	15	US-10-369-493-17433	Sequence 17433, A
18	1402.5	65.1	423	15	US-10-369-493-10557	Sequence 10557, A
19	1402.5	65.1	426	15	US-10-369-493-17886	Sequence 17886, A
20	1400	65.0	424	15	US-10-369-493-11819	Sequence 11819, A
21	1396.5	64.8	431	15	US-10-282-122A-52760	Sequence 52760, A
22	1390	64.5	420	15	US-10-369-493-14739	Sequence 14739, A
23	1390	64.5	423	15	US-10-369-493-15222	Sequence 15222, A
24	1381.5	64.1	430	15	US-10-369-493-23299	Sequence 23299, A
25	1377	63.9	431	15	US-10-282-122A-46167	Sequence 46167, A
26	1370.5	63.6	448	15	US-10-282-122A-52925	Sequence 52925, A
27	1366.5	63.4	427	15	US-10-369-493-21006	Sequence 21006, A
28	1361.5	63.2	430	15	US-10-369-493-19491	Sequence 19491, A
29	1360.5	63.1	430	15	US-10-282-122A-60954	Sequence 60954, A
30	1358.5	63.0	426	15	US-10-369-493-21730	Sequence 21730, A
31	1358.5	63.0	431	15	US-10-282-122A-51920	Sequence 51920, A
32	1352.5	62.8	427	15	US-10-282-122A-47804	Sequence 47804, A
33	1351.5	62.7	417	15	US-10-369-493-9097	Sequence 9097, Ap
34	1348.5	62.6	426	15	US-10-369-493-4688	Sequence 4688, Ap
35	1348.5	62.6	427	15	US-10-369-493-7447	Sequence 7447, Ap
36	1348.5	62.6	427	15	US-10-282-122A-49282	Sequence 49282, A
37	1346.5	62.5	428	16	US-10-375-266-38	Sequence 38, Appl
38	1346	62.5	429	9	US-09-815-243-5012	Sequence 5012, Ap
39	1346	62.5	432	9	US-09-815-243-10854	Sequence 10854, A
40	1346	62.5	432	15	US-10-282-122A-42554	Sequence 42554, A
41	1345.5	62.4	424	15	US-10-369-493-7762	Sequence 7762, Ap
42	1336.5	62.0	428	15	US-10-369-493-9970	Sequence 9970, Ap
43	1335.5	62.0	426	15	US-10-369-493-32	Sequence 32, Appl
44	1327.5	61.6	424	15	US-10-369-493-20662	Sequence 20662, A
45	1326	61.5	422	15	US-10-369-493-23433	Sequence 23433, A

Qy	1	VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLDDGDRYL	60
Db	1	VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLDDGDRYL	60
Qy	61	KGVLKAVENVNEEIGDELAGEADDQRLIDAMIKLDTANKSRIGANAILGVSNVAKA	120
Db	61	KGVLKAVENVNEEIGDELAGEADDQRLIDAMIKLDTANKSRIGANAILGVSNVAKA	120

Query Match 100.0%; Score 2155; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 2.6e-171;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AADGAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFPIAGETPSEALRN 180
 DB 121 AADGAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFPIAGETPSEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDGGFAPSGVSTREALDLIVEAEKAGFTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDGGFAPSGVSTREALDLIVEAEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDADVMAHRAGYTSM 360
 DB 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDADVMAHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFOG 425
 DB 421 PRFOG 425
 RESULT 2
 US-10-728-947-2
 ; Sequence 2, Application US/10728947
 ; Publication No. US20040203941
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockett, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jörn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10/728,947
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-728-947-2
 Query Match 100.0%; Score 2155; DB 16; Length 425;
 Best Local Similarity 100.0%; Pred. No. 2.6e-171;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTPTAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 DB 1 VAEIMHVFAREILDSRGNTPTAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 QY 61 KGVLKAVENVNBEIGDELGLADDDQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
 DB 61 KGVLKAVENVNBEIGDELGLADDDQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
 QY 121 AADGAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFPIAGETPSEALRN 180
 DB 121 AADGAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFPIAGETPSEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDGGFAPSGVSTREALDLIVEAEKAGFTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDGGFAPSGVSTREALDLIVEAEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300

DB 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDADVMAHRAGYTSM 360
 DB 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDADVMAHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFOG 425
 DB 421 PRFOG 425
 RESULT 3
 US-09-738-626-4585
 ; Sequence 4585, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIALI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 4585
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4585
 Query Match 99.9%; Score 2152; DB 9; Length 425;
 Best Local Similarity 99.8%; Pred. No. 4.7e-171;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTPTAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 DB 1 VAEIMHVFAREILDSRGNTPTAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 QY 61 KGVLKAVENVNBEIGDELGLADDDQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
 DB 61 KGVLKAVENVNBEIGDELGLADDDQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
 QY 121 AADGAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFPIAGETPSEALRN 180
 DB 121 AADGAGLPLFRYIGGPNHVLVPMNIIINGGAHADSGVDVQEFPIAGETPSEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLDGGFAPSGVSTREALDLIVEAEKAGFTPGKDIALAL 240
 DB 181 GAEVYHALKSVIKKGLSTGLDGGFAPSGVSTREALDLIVEAEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 QY 301 GDKQVIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDADVMAHRAGYTSM 360

Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFQ 425
 Db 421 PRFQ 425

RESULT 4
 US-09-860-768-4
 ; Sequence 4, Application US/09860768
 ; Patent No. US20020082403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pöhler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/09/860,768
 ; CURRENT FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-860-768-4

Query Match 99.8%; Score 2151; DB 9; Length 425;
 Best Local Similarity 99.8%; Pred. No. 5.6e-171;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 Db 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 QY 61 KGVLKAVENNEEIGDELGLAGLEADDORLIDEAMIKLDGTANKSRILGANAILGVSM 120
 Db 61 KGVLKAVENNEEIGDELGLAGLEADDORLIDEAMIKLDGTANKSRILGANAILGVSM 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITGAETTFSEALRN 180
 Db 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITGAETTFSEALRN 180
 QY 181 GAEVTHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALAL 240
 Db 181 GAEVTHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEQGHSAEMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHFEQGHSAEMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFQ 425
 Db 421 PRFQ 425

RESULT 5

US-10-728-947-4
 ; Sequence 4, Application US/10728947
 ; Publication No. US20040220394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pöhler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10/728,947
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-728-947-4

Query Match 99.8%; Score 2151; DB 16; Length 425;
 Best Local Similarity 99.8%; Pred. No. 5.6e-171;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 Db 1 VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHELDDGDRYL 60
 QY 61 KGVLKAVENNEEIGDELGLAGLEADDORLIDEAMIKLDGTANKSRILGANAILGVSM 120
 Db 61 KGVLKAVENNEEIGDELGLAGLEADDORLIDEAMIKLDGTANKSRILGANAILGVSM 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITGAETTFSEALRN 180
 Db 121 AADSAGLPLFRYIGGPNHVLVPPMNIINGGAHDSGVVDVQEFMIAPITGAETTFSEALRN 180
 QY 181 GAEVTHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALAL 240
 Db 181 GAEVTHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVEAIEKAGTTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEQGHSAEMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHFEQGHSAEMANYAELVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVYAGRSAP 420
 QY 421 PRFQ 425
 Db 421 PRFQ 425

RESULT 6
 US-10-781-014-72
 ; Sequence 72, Application US/10781014
 ; Publication No. US20040180408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pompejus, Markus
 ; APPLICANT: Kröger, Burkhard
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Haberhauser, Gregor
 ; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN CARBON METABOLISM AND ENERGY

; TITLE OF INVENTION: PRODUCTION
 ; FILE REFERENCE: BGI-126CPCN
 ; CURRENT APPLICATION NUMBER: US/10/781.014
 ; CURRENT FILING DATE: 2004-02-17
 ; PRIOR APPLICATION NUMBER: US 09/602,740
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/141,031
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: 60/143,208
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 60/151,572
 ; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: DE 19931412.8
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931413.6
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931419.5
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931420.9
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931424.1
 ; PRIOR FILING DATE: 1999-07-08
 ; PRIOR APPLICATION NUMBER: DE 19931428.4
 ; PRIOR FILING DATE: 1999-07-08
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 784
 ; SEQ ID NO 72
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-781-014-72

Query Match 99.7%; Score 2149; DB 16; Length 425;
 Best Local Similarity 99.8%; Pred. No. 8.3e-171;
 Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGDRYLG 60
 QY 61 KGVLKAVENVNEEIGDELADDDORLIDEAMIKLDTANKSRKLGANAILGVSMVAKA 120
 DB 61 KGVLKAVENVNEEIGDELADDDORLIDEAMIKLDTANKSRKLGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIAPIGAEFTSEALRN 180
 DB 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIAPIGAEFTSEALRN 180
 QY 181 GAENVHAKSVIKKGLSTGLDGGPAPSVGSGTREALDLIVEAIEKAGFTPGKDIALAL 240
 DB 181 GAENVHAKSVIKKGLSTGLDGGPAPSVGSGTREALDLIVEAIEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEYTNLTATI 300
 QY 301 GDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 DB 301 GDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 QY 421 PRFQG 425
 DB 421 PRFQG 425

RESULT 7
 US-10-494-836-8
 ; Sequence 8, Application US/10494836
 ; Publication No. US20050014233A1

; GENERAL INFORMATION:
 ; APPLICANT: Zelder, Oskar
 ; APPLICANT: Pompejus, Markus
 ; APPLICANT: Schroder, Hartwig
 ; APPLICANT: Kroger, Burkhard
 ; APPLICANT: Klopprogge, Corinna
 ; APPLICANT: Haberhauser, Gregor
 ; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy produc
 ; FILE REFERENCE: BGI-167US
 ; CURRENT APPLICATION NUMBER: US/10/494,836
 ; CURRENT FILING DATE: 2004-05-05
 ; PRIOR APPLICATION NUMBER: PCT/EP02/12135
 ; PRIOR FILING DATE: 2002-10-31
 ; PRIOR APPLICATION NUMBER: DE 101 54 270.4
 ; PRIOR FILING DATE: 2001-11-05
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 8
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-494-836-8

Query Match 99.7%; Score 2149; DB 17; Length 425;
 Best Local Similarity 99.8%; Pred. No. 8.3e-171;
 Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGDRYLG 60
 QY 61 KGVLKAVENVNEEIGDELADDDORLIDEAMIKLDTANKSRKLGANAILGVSMVAKA 120
 DB 61 KGVLKAVENVNEEIGDELADDDORLIDEAMIKLDTANKSRKLGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIAPIGAEFTSEALRN 180
 DB 121 AADSAGLPLFRYIGGNPAHVLVPVMNIIIGGAHADSGVDVQEFMIAPIGAEFTSEALRN 180
 QY 181 GAENVHAKSVIKKGLSTGLDGGPAPSVGSGTREALDLIVEAIEKAGFTPGKDIALAL 240
 DB 181 GAENVHAKSVIKKGLSTGLDGGPAPSVGSGTREALDLIVEAIEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEYTNLTATI 300
 DB 241 DVASSEFFKDGTYHFEGGQHSAAEMANYAELVDAYPIVSIEDPLQEDDWEYTNLTATI 300
 QY 301 GDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 DB 301 GDKVQIVGDDPFVTPNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHARAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 DB 361 SHRSGETEDTTIADLAVALNCGQIKTGAPASDRVAKYNQLLRIEQLLDGAGVYAGRSF 420
 QY 421 PRFQG 425
 DB 421 PRFQG 425

RESULT 8
 US-10-282-122A-53908
 ; Sequence 53908, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant


```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 62133
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62133

```

```

Query Match 72.3%; Score 1558.5; DB 15; Length 429;
Best Local Similarity 72.9%; Pred. No. 1.9e-121;
Matches 307; Conservative 46; Mismatches 67; Indels 1; Gaps 1;

Qy 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLKGV 63
Db 4 IQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLKGV 63

Qy 64 LKAVENVAEEIGDELGLAGLEADQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAKAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGPKSRKLGANAILGVSLAVAKAAD 123

Qy 124 SAGLPLFRYIGGNPAHVLVPMNMNINGGAHADSGVDVQEFMIAPIGAEFTSEALRNQAE 183
Db 124 SAEPLFRYIGGNPAHVLVPMNMNINGGAHADTAVIDIQEFMVAIPGAPSAEALRWGAE 183

Qy 184 VYHALKSIVKEKGLSTGLDGGEGFAPSVGSTREALDLIVEAEKAGFTPGKDIALLDVA 243
Db 184 VYHSLSKIVKEKGLSTGLDGGEGFADVPAGTAAALDLIGRAIESAGFKLGTVDVALDAA 243

Qy 244 SSEFFKDTG-YHFEQGQHSAAEMANVYAEVLVDAYPIVSIEDPQDDEWEGYTNLTATIGD 302
Db 244 ATEFYSDGTGYKFEQSTABQMAEFYAGLLGAYPLVSIEDPQDDEWEGYTNLTATIGD 303

Qy 303 KQVIVGDDFFVTNPERLKEGIAKKAANSILVKNOIGTLTETTFDAVDMHARAGYTSMMSH 362
Db 304 RVQLVGGDDVFTNPERLKEGIEKGVANALLVKVNOIGTLTETTFDAVALAHSHGYRTMMSH 363

Qy 363 RSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLDAGVYAGRSAPPR 422
Db 364 RSGETEDTTIADLAVAGSGQIKTGAPARSEKVKYNQLLRIEALGDAARYAGDLAPPR 423

Qy 423 F 423
Db 424 F 424

```

```

RESULT 11
US-10-156-761-11070
; Sequence 11070, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11070
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11070

```

```

Query Match 72.2%; Score 1555.5; DB 14; Length 428;
Best Local Similarity 73.2%; Pred. No. 3.4e-121;
Matches 311; Conservative 42; Mismatches 71; Indels 1; Gaps 1;

Qy 1 VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 59
Db 1 VPSIDVVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVFAELRGGDRYQ 60

Qy 60 GKGVLKAVENVAEEIGDELGLAGLEADQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAK 119
Db 61 GKGVEKAVLAVIEQIGPELVGYDATEQRLIDQAMFDDATDNKSGSLGANAILGVSLAVAH 120

Qy 120 AAADSAGLPLFRYIGGNPAHVLVPMNMNINGGAHADSGVDVQEFMIAPIGAEFTSEALR 179
Db 121 AASEASDPLFRYIGGNPAHVLVPMNMNINGGSHADSNVDIQEFMIAPIGAEFTSEALR 180

Qy 180 NGAEVYHALKSIVKEKGLSTGLDGGEGFAPSVGSTREALDLIVEAEKAGFTPGKDIALLA 239
Db 181 WGAEVYHTLAKVLKTKGLSTGLDGGEGFAPNLESNRAALDLIIIEAKQAGYIFGEQIALA 240

Qy 240 LDVASSFFKDTGTHFEQGQHSAAEMANVYAEVLVDAYPIVSIEDPQDDEWEGYTNLTAT 299
Db 241 LDVAASEFYKDTGTHFEQGQHSAAEMTEYEEELVSAVPLVSIEDPQDDEWAGKVIKTK 300

Qy 300 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKNOIGTLTETTFDAVDMHARAGYTSM 359
Db 301 LGDKVQIVGDDLFPVTNPERLARGIEGSAVALLVKNQIGSLTETTFDAVLAQRNGPKCM 360

Qy 360 MSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRIEQLDAGVYAGRSA 419
Db 361 MSHRSGETEDTTIADLAVAVNCQIKTGAPARSDRVAKYNQLLRIEILLDDAAEYAGRSA 420

Qy 420 FPRQ 424
Db 421 FPRR 425

```

```

RESULT 12
US-10-282-122A-64490
; Sequence 64490, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos

```

; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR FILING DATE: 2001-02-16
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 64490
 ; LENGTH: 429
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-10-282-122A-64490

 Query Match 72.0%; Score 1551.5; DB 15; Length 429;
 Best Local Similarity 73.2%; Pred. No. 7.3e-121;
 Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

 QY 4 IMHVFAEILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
 DB 4 IEQVFAEILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
 QY 64 LKAVENVEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAAD 123
 DB 64 QKAVQAVLDIGPAVIGLNADQRLVDQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 123
 QY 124 SAGLPLEFYGGPNNAHILPVPMMNININGGAHDSGVQVEFMIAPIGAETPSFSEALRNGAE 183
 DB 124 SAGLPLEFYGGPNNAHILPVPMMNININGGAHDSGVQVEFMIAPIGAETPSFSEALRNGAE 183
 QY 184 VTHALSKVKEGLSTGLDEGGPAPSVGSTREALDLIVEATEKAGFTPGKDIALALDVA 243
 DB 184 VTHALSKVKEGLSTGLDEGGPAPSVGSTREALDLIVEATEKAGFTPGKDIALALDVA 243
 QY 244 SFEFFKDTG-YHFEQGHSAARMANVYAEVDAPIVSIEDPLQBDDEWEGYTNLTATGD 302
 DB 244 ATEFTTDTGTGVFEQTRTADQMTFYGALLGAYPLVSTEDPLSDDDWDGWAALTASIGD 303
 QY 303 KVOIQDFFVTNPRELKEGIAKKAANSILVKVNOIGTITETFDVADNHAHAGYTSWMSH 362
 DB 304 RVQIQDFFVTNPRELKEGIERGVANALLVKVNOIGTITETFDVADNHAHAGYTSWMSH 363
 QY 363 RSGETEDTMIADLAVALCGQIKTGAPARSDRVAKYNOLLRIEALGAGVYAGRSAPPR 422

Db 364 RSGETEDTMIADLAVALCGQIKTGAPARSDRVAKYNOLLRIEALGAGVYAGRSAPPR 422
 QY 423 F 423
 Db 424 F 424

 RESULT 13
 US-10-282-122A-62696
 ; Sequence 62696, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR FILING DATE: 2001-02-16
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 62696
 ; LENGTH: 439
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium bovis
 US-10-282-122A-62696

 Query Match 72.0%; Score 1551.5; DB 15; Length 439;
 Best Local Similarity 73.2%; Pred. No. 7.6e-121;
 Matches 308; Conservative 41; Mismatches 71; Indels 1; Gaps 1;

 QY 4 IMHVFAEILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 63
 DB 14 IEQVFAEILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLGKV 73
 QY 64 LKAVENVEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKAAAD 123
 DB 74 QKAVQAVLDIGPAVIGLNADQRLVDQALVDLDTGTPDKSRGLGNAILGVSLAVAKAAAD 133
 QY 124 SAGLPLEFYGGPNNAHILPVPMMNININGGAHDSGVQVEFMIAPIGAETPSFSEALRNGAE 183
 DB 134 SAGLPLEFYGGPNNAHILPVPMMNININGGAHDSGVQVEFMIAPIGAETPSFSEALRNGAE 193


```

Db 124 ASGLPLYRYVGGPNAHLFPVPMNIIINGGAHADNPIDPOEFMILPVGAENIREAVRMGSE 183
Qy 184 VYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALALDVA 243
Db 184 VFHTLKKELSAQGHNTNVGDEGGFAPGLESAPALDFIMKSIKAGTRGPEDMYVGLDCA 243
Qy 244 SSEFFKDGTYHPGGOHS--AAEMANVYAELVDAYPIVSIEDPLQEDDWEGYTNLTATIG 301
Db 244 STEFFKDGKYVLEGEGRITLEPGAMAEYLAELVNKYPIISVEDGMAEDDWEGWKTLTDLVG 303
Qy 302 DKVOIVGDDFFVTNPERLKEGIKKAANSILVKVNOIGTITETFDVADNAHRAGYTSMS 361
Db 304 NKQQLVGDDLFTVNSARLDRGIRMGVANSILVKVNOIGSLSETLDAVETAHKAGYTAVMS 363
Qy 362 HRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEOLLGDAGVYAGRS 418
Db 364 HRSGETEDTTIADLAVATNCGOIKTGLARSRLAKYNOLIRIEMLGPOAYAGRS 420

```

Search completed: September 29, 2005, 14:45:39
 Job time : 1028 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:02:19 ; Search time 6707 Seconds
(without alignments)
11400.377 Million cell updates/sec

Title: US-10-728-947-3

Perfect score: 1578

Sequence: 1 ggcctggggataggtagttt.....ctcaagcagggagcgtgctt 1578

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6	AR490886
2	1576.4	99.9	1578	6	AR490885 Sequence
3	1576.4	99.9	1578	6	AX136862 Sequence
4	1576.4	99.9	1578	6	BD014844 Polynucle
5	1576.4	99.9	333150	1	AP005277 Coryneb
6	1576.4	99.9	348475	1	AX927150 Coryneb
7	1576.4	99.9	349980	6	AX127146 Coryneb
8	1401.8	88.8	1405	6	AX763190 Sequence
9	1394.8	88.4	1398	6	AX064945 Sequence
10	1273.4	80.7	1275	6	BD163286 Novel pol
11	1273.4	80.7	1275	6	AX121169 Sequence
12	1034.8	65.6	300750	1	AP005217 Coryneb
13	915.8	58.0	347625	1	AX248356 Coryneb
14	732.4	46.4	276800	1	SC0939115 Streptomy
15	700	44.4	110000	1	AE016822_17 Continuation (18 o
16	696.8	44.2	296300	1	AP005035 Streptomy
17	680.6	43.1	110000	1	AP006618_50 Continuation (51 o
18	657.6	41.7	303855	1	AE017230 Mycobacte
19	640.8	40.6	30159	6	CQ363776 Sequence

20 640.8 40.6 110000 1 AE017283_05 Continuation (6 of
21 632.6 40.1 45191 6 AX704278 AX704278 Sequence
22 632.6 40.1 110000 1 AE000516_11 Continuation (12 o
23 632.6 40.1 327650 1 BX248337 Mycobacte
24 632.6 40.1 349306 1 BX842575 Mycobacte
25 590 37.4 329100 1 SME591787 ALS91787 Sinorhizo
26 583.4 37.0 344050 1 MLEPRTN1 AE014725 Bifidobac
27 583 36.9 10029 1 AE014725 Bifidobac
28 583 36.9 349980 6 AX492784 AE014725 Bifidobac
29 583 36.9 349980 6 AX492784 AE014725 Bifidobac
30 583 36.9 349980 6 AX492784 AE014725 Bifidobac
31 583 36.9 349980 6 AX492784 AE014725 Bifidobac
32 579.4 36.7 12058 1 AE009103 AE009103 Sequence
33 579.4 36.7 12407 1 AE008068 AE008068 Agrobacte
34 578 36.6 10934 1 AE005846 AE005846 Caulobact
35 575.4 36.5 347660 1 AP002994 AP002994 Mesorhizo
36 574.8 36.4 12709 1 AE014414 AE014414 Brucella
37 571.6 36.2 13776 1 AE009526 AE009526 Brucella
38 570.2 36.1 1290 1 AX181039 BX572602 Rhodopseu
39 569.8 36.1 349981 1 BX572602 BX572602 Rhodopseu
40 559.6 35.5 300600 1 AP005952 AP005952 Bradyrhiz
41 557.8 35.3 193050 1 AL646062 AL646062 Bradyrhiz
42 555.6 35.2 197615 5 BX470161 BX470161 Zebraphish
43 550.2 34.9 110000 1 AE017180_24 Continuation (25 o
44 550.2 34.9 110000 1 AE017180_25 Continuation (26 o
45 542.8 34.4 110000 1 BX571965_27 Continuation (28 o

ALIGNMENTS

RESULT 1
AR490886
LOCUS AR490886 1578 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 3 from patent US 6713289.
ACCESSION AR490886
VERSION AR490886.1 GI:47258405
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1578)
AUTHORS Mockel,B., Pfefferle,W., Hermann,T., Puhler,A., Kalinowski,J. and Bathe,B.
TITLE Nucleotide sequences which code for the eno gene
JOURNAL Patent: US 6713289-A 3 30-MAR-2004;
FEATURES
source location/Qualifiers
1..1578
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1578; DB 6; Length 1578;
Best Local Similarity 100.0%; Pred. No. 3.4e-289;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCTGGGATATGGTAGTTTCGCCACTAATTTCACTGATTCCTCATCGAACACAGA 60
Db 1 GCCTGGGATATGGTAGTTTCGCCACTAATTTCACTGATTCCTCATCGAACACAGA 60
Qy 61 TTCGTGCAACAATTTGGGTGTAGAGTGTGAAGACATTTTCATCAGCGTAATATTCCTAG 120
Db 61 TTCGTGCAACAATTTGGGTGTAGAGTGTGAAGACATTTTCATCAGCGTAATATTCCTAG 120
Qy 121 TTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 180
Db 121 TTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 180
Qy 181 GAAATTTCCGACTCCCGGGTAAACCCACCGTCGAGCAGAGGTTTTCCTGGATGACGGT 240
Db 181 GAAATTTCCGACTCCCGGGTAAACCCACCGTCGAGCAGAGGTTTTCCTGGATGACGGT 240
Qy 241 TCCACCGGTGTGGCAGGTGTTTCCATCCGCGGCATCCACCGCGCTCATGAG 300

Db 421 GAAGCAATGATCAAGCTTGTATGACACCGCCAAAGTCCCGCTGGTGCACCAATC 480
Qy 481 CTTGTGTTTCCATGGCTGTGCAAAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTC 540
Db 481 CTTGTGTTTCCATGGCTGTGCAAAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTC 540
Qy 541 CGTACATCGGTGAGCAAAACGACACGCTTCCAGTTCCTCAATGATGAACATCATCAAC 600
Db 541 CGTACATCGGTGAGCAAAACGACACGCTTCCAGTTCCTCAATGATGAACATCATCAAC 600
Qy 601 GTGGGGCTCAGCTGACTCCGGTGTGAGCTTACGGAATTCATGATCGTCCATCGGT 660
Db 601 GTGGGGCTCAGCTGACTCCGGTGTGAGCTTACGGAATTCATGATCGTCCATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCTGAAGTCC 720
Qy 721 GTCATCAAGAAAGGCTGTCCACCGACTTGGCGATGAGGGCGCTTCGCTCCTCC 780
Db 721 GTCATCAAGAAAGGCTGTCCACCGACTTGGCGATGAGGGCGCTTCGCTCCTCC 780
Qy 781 GTGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGCTTC 840
Db 781 GTGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGCTTC 840
Qy 841 ACCCAGGCAAGACATCGCTCTGCTGTGAGCTTCTCTGAGTTCCTCAAGGAC 900
Db 841 ACCCAGGCAAGACATCGCTCTGCTGTGAGCTTCTCTGAGTTCCTCAAGGAC 900
Qy 901 GGCACCTACCACTTCCGAGGTGCGCAGCACTCCGACGCTGAGATGCAAACTTTACGCT 960
Db 901 GGCACCTACCACTTCCGAGGTGCGCAGCACTCCGACGCTGAGATGCAAACTTTACGCT 960
Qy 961 GAGCTGTTGACGCGTACCAATCGTCTCCATCGAGGACCCACTGCGAGGATGACTGG 1020
Db 961 GAGCTGTTGACGCGTACCAATCGTCTCCATCGAGGACCCACTGCGAGGATGACTGG 1020
Qy 1021 GAGGTTACCAACCTCACCAGCAACCTACCGGACCAAGGTTTCAGATGTCGCGACAC 1080
Db 1021 GAGGTTACCAACCTCACCAGCAACCTACCGGACCAAGGTTTCAGATGTCGCGACAC 1080
Qy 1081 TTCTTCGTCAACACCTGAGCGCTGAGAGGCGCATCGCTAAGAAGGCTCCCACTCC 1140
Db 1081 TTCTTCGTCAACACCTGAGCGCTGAGAGGCGCATCGCTAAGAAGGCTCCCACTCC 1140
Qy 1141 ATCTGTTTAAAGTGAAACAGATCGGTACCTCACCAGACCTTCGACGCTTCGACATG 1200
Db 1141 ATCTGTTTAAAGTGAAACAGATCGGTACCTCACCAGACCTTCGACGCTTCGACATG 1200
Qy 1201 GCTCACCGGAGGCTTACCTCCATGATGTCCACCGTTCGGTTCGAGACGACAC 1260
Db 1201 GCTCACCGGAGGCTTACCTCCATGATGTCCACCGTTCGGTTCGAGACGACAC 1260
Qy 1261 ACCATTGCTGACTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGTGCTCCAGCA 1320
Db 1261 ACCATTGCTGACTCGCAGTTGCACTCAACTGTGCGCAGATCAAGACTGTGCTCCAGCA 1320
Qy 1321 CGTCCGACCGGTGTCGCAAGTACCAACGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTCCGACCGGTGTCGCAAGTACCAACGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCGGCGTCTACGAGGTGCGAGCGCATCCACGCTTTCAGGGCTTAATAAAGCGCTT 1440
Db 1381 GCGGCGTCTACGAGGTGCGAGCGCATCCACGCTTTCAGGGCTTAATAAAGCGCTT 1440
Qy 1441 TTCAGCGCGGTACCTCAAGTTGCGGGGCTGTGCTTACTTACTTACTTACTTACTT 1500
Db 1441 TTCAGCGCGGTACCTCAAGTTGCGGGGCTGTGCTTACTTACTTACTTACTTACTT 1500
Qy 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAACTCATAAAGGCTTGTTCCTGTCT 1560

Db 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAACTCATAAAGGCTTGTTCCTGTCT 1560
Qy 1561 CAAGCAGGCAACGTGCTT 1578
Db 1561 CAAGCAGGCAACGTGCTT 1578
RESULT 3
AX136862
LOCUS AX136862 1578 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1090998.
ACCESSION AX136862
VERSION AX136862.1 GI:14273213
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Moeckel, B., Pfefferle, W., Hermann, T., Puehler, A., Kalinowski, J. and Bathe, B.
TITLE Nucleotide sequences coding for the eno gene
JOURNAL Patent: EP 1090998-A 1 11-APR-2001;
Degussa AG (DE)
FEATURES
Location/Qualifiers
source 1. 1578
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
151..1428
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC39860.1"
/db_xref="GI:14273213"
CDS
Query Match 99.9%; Score 1576.4; DB 6; Length 1578;
Best Local Similarity 99.9%; Pred. No. 6.8e-289;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCACTGATTCGCTCATCGAAACAAGA 60
Db 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCACTGATTCGCTCATCGAAACAAGA 60
Qy 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTCATCAGTGAATAATTTCTAG 120
Db 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTCATCAGTGAATAATTTCTAG 120
Qy 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGATTCGCTCGC 180
Qy 181 GAATTTCTCGACTCCCGGCTTAACCCAGCTCGAGGACAGAGTTTTCCTGGATGACGGT 240
Db 181 GAATTTCTCGACTCCCGGCTTAACCCAGCTCGAGGACAGAGTTTTCCTGGATGACGGT 240
Qy 241 TCCCAACCGGTGTGCGAGGTGTTCATCCGCGGCATCCACCGGCTCCACGAGGCTCATGAG 300
Db 241 TCCCAACCGGTGTGCGAGGTGTTCATCCGCGGCATCCACCGGCTCCACGAGGCTCATGAG 300
Qy 301 CTGCGTCAACGCTGCGATTCGCTACCTCGGCAAGGGCGTTTGAAGGAGTTGAAAGCGTC 360
Db 301 CTGCGTCAACGCTGCGATTCGCTACCTCGGCAAGGGCGTTTGAAGGAGTTGAAAGCGTC 360

Qy	361	AACGAAAGAAATTCGGCGACGAGCTCGCTGGCCCTAGAGGCTGACGATACGCGCCTCATCGAC	420
Db	361	AACGAAAGAAATTCGGCGACGAGCTCGCTGGCCCTAGAGGCTGACGATACGCGCCTCATCGAC	420
Qy	421	GAAAGCAATGATCAAGCTTTGATGGCACCAGCAACGAGTCCCGCTGGGGTGCAGAACGCAATC	480
Db	421	GAAAGCAATGATCAAGCTTTGATGGCACCAGCAACGAGTCCCGCTGGGGTGCAGAACGCAATC	480
Qy	481	CTTGGTGTTTCCATAGGCTGTTGCAAAAGGCTGCTGCTGATTCGCGACAGCCCTCCCACTGCTTC	540
Db	481	CTTGGTGTTTCCATAGGCTGTTGCAAAAGGCTGCTGCTGATTCGCGACAGCCCTCCCACTGCTTC	540
Qy	541	CGCTACATCGGTGACCAAAACGCAACGCTTCTCCAGTTCCAAATGATGAACATCATCAAC	600
Db	541	CGCTACATCGGTGACCAAAACGCAACGCTTCTCCAGTTCCAAATGATGAACATCATCAAC	600
Qy	601	GGTGGCGCTCAGCGTGAATCTCGGTGTTGACGTTCAGGAAATTCATGATGCTGCATCGGT	660
Db	601	GGTGGCGCTCAGCGTGAATCTCGGTGTTGACGTTTACGAAATTCATGATGCTGCATCGGT	660
Qy	661	GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACACGACACTGAAGTCC	720
Db	661	GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACACGACACTGAAGTCC	720
Qy	721	GTCAATCAAGGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGCGGCTTGCTCTCTTC	780
Db	721	GTCAATCAAGGAAAGGGCCTGTCCACCGGACTTGGCGATGAGGCGGCTTGCTCTCTTC	780
Qy	781	GTCCGCTCCACCGTGAGGCTCTTGAACCTTATCGTTAAGGCAATCGAAGGCTGGCTTC	840
Db	781	GTCCGCTCCACCGTGAGGCTCTTGAACCTTATCGTTAAGGCAATCGAAGGCTGGCTTC	840
Qy	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGACGCTGCTTCTCTAGTTTCTTCAAGGAC	900
Db	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGACGCTGCTTCTCTAGTTTCTTCAAGGAC	900
Qy	901	GGCACTACCACTTCGAAGGTGGCGACGACATCCCGCAGCTGAGATGGCAAACTGTTACGCT	960
Db	901	GGCACTACCACTTCGAAGGTGGCGACGACATCCCGCAGCTGAGATGGCAAACTGTTACGCT	960
Qy	961	GAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGGACCCACCTGCGAGGAAGATGACTGG	1020
Db	961	GAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGGACCCACCTGCGAGGAAGATGACTGG	1020
Qy	1021	GAGGTTTACACCAACTCACCAGCAACATCGCGGACCAAGTTCAGATCGTTGGCGACGAC	1080
Db	1021	GAGGTTTACACCAACTCACCAGCAACATCGCGGACCAAGTTCAGATCGTTGGCGACGAC	1080
Qy	1081	TTCTTTCGTCCACCAACCTTGAGCGCTGGAAGGAGGGCATCGCTAAGAAGGCTGCCAACTCC	1140
Db	1081	TTCTTTCGTCCACCAACCTTGAGCGCTGGAAGGAGGGCATCGCTAAGAAGGCTGCCAACTCC	1140
Qy	1141	ATCTCGTTTAAGGTGAACCAAGATCGGTACCTCAACCGAGACCTTTCGACGCTGTGCAATG	1200
Db	1141	ATCTCGTTTAAGGTGAACCAAGATCGGTACCTCAACCGAGACCTTTCGACGCTGTGCAATG	1200
Qy	1201	GCTCACCAGCGAGGCTACACTCCATGATGTCCACCGTTCGGGTGAGACCGAGGACCC	1260
Db	1201	GCTCACCAGCGAGGCTACACTCCATGATGTCCACCGTTCGGGTGAGACCGAGGACCC	1260
Qy	1261	ACCAATGCTGACCTCGCAGTTTGCACTCAAATGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Db	1261	ACCAATGCTGACCTCGCAGTTTGCACTCAAATGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Qy	1321	CGTTCCGACCGGTGCGCAAAAGTACAAACAGCTTCTTCGCGATCGAGCAGCTGCTTGGCGAC	1380
Db	1321	CGTTCCGACCGGTGCGCAAAAGTACAAACAGCTTCTTCGCGATCGAGCAGCTGCTTGGCGAC	1380
Qy	1381	GCCGGCGTCTACGAGGTGCGAGCGCATTTCCACGCTTTCAGGGCTAAATAAAGACGCTT	1440
Db	1381	GCCGGCGTCTACGAGGTGCGAGCGCATTTCCACGCTTTCAGGGCTAAATAAAGACGCTT	1440
Qy	1441	TTCGACGCCCGGTAACTCAAGGTTGCGCGCGCTGTTTGCCTTACTACTGTTACTGCTGT	1500

Db	1441	TTTGACCCCGGTAACTCAAGCTTGCCTGCGGCGTGTGCTTACTACTGTTACTGCTGT	1500
Qy	1501	GACTATGATCGAGGATTATGCGAAGCAGAGAAAGAACTCATAAAGGCTTGTTCCTGCTCT	1560
Db	1501	GACTATGATCGAGGATTATGCGAAGCAGAGAAAGAACTCATAAAGGCTTGTTCCTGCTCT	1560
Qy	1561	CAAGCAGGGAACGTGCTT	1578
Db	1561	CAAGCAGGGAACGTGCTT	1578
RESULT 4			
BD014844			
LOCUS			
DEFINITION	BD014844	1578 bp DNA linear	PAT 27-AUG-2002
ACCESSION	BD014844.1	GI:22555651	
VERSION	JP 2001161380-A/1.		
KEYWORDS	Corynebacterium glutamicum		
SOURCE	Corynebacterium glutamicum		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1 (bases 1 to 1578)		
AUTHORS	Meckel, V., Pfeifferle, W., Hermann, T., Puehler, A., Kalinowski, J. and Baate, B.		
TITLE	Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid		
JOURNAL	Patent: JP 2001161380-A 1 19-JUN-2001;		
COMMENT	DEGUSA HUELS AG OS Corynebacterium glutamicum PN JP 2001161380-A/1 PD 19-JUN-2001 PF 04-OCT-2000 JP 2000305110 PI 05-OCT-1999 DE 19947791.4 PR VETTYNA MECKEL, WALTER PFEIFFERLE, THOMAS HERMANN, ALFRED PUEHLER, PI JOERN KALINOWSKI, BRIGITTE BAATE PC C12N15/09, C12N1/21, C12N9/88, C12P13/04, C12P13/08//C12N15/09, PC C12R1:15), PC (C12P13/04, C12R1:15), (C12P13/08, C12R1:15), C12N15/00, (C12N15/00, PC C12R1:15) CC Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid		
FEATURES			
source	1..1578	/organism="Corynebacterium glutamicum"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:1718"	
ORIGIN			
Query Match	99.98;	Score 1576.4;	DB 6; Length 1578;
Best Local Similarity	99.98;	Pred. No. 6.8e-289;	
Matches 1577;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	GGCTGGGGATATGGGTAGTTTCGCCACTAATTTCCACTGATTCGCTTCATCGAAACAAGA	60
Db	1	GGCTGGGGATATGGGTAGTTTCGCCACTAATTTCCACTGATTCGCTTCATCGAAACAAGA	60
Qy	61	TTCTGTGCAACAAATTTGGGTGTAGCGTGAATTTGAAGACATTTGATCAGCTGAATATTTCTAG	120
Db	61	TTCTGTGCAACAAATTTGGGTGTAGCGTGAATTTGAAGACATTTGATCAGCTGAATATTTCTAG	120
Qy	121	TTAGCTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGACGATTTTCGCTCGC	180
Db	121	TTAGCTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAAATCATGACGATTTTCGCTCGC	180
Qy	181	GAATTTCTCGACTCCCGGGTAAACCCAAACCGTTCGAGGCAGAGGTTTTCCTCGATGACGGT	240

DB	181	TTCTCGACTCCGCGGGTAACCAACCGTCGAGGCGAGGTTTTCTCTGATCAGCGT	240
QY	241	TCCCAACGGTGTGCGAGGTGTTCCATCCGCGCGCATCCACCGGCGTCCACGAGGCTCATGAG	300
DB	241	TCCCAACGGTGTGCGAGGTGTTCCATCCGCGCGCATCCACCGGCGTCCACGAGGCTCATGAG	300
QY	301	CTGCGTGAACGGTGGCGATGCTCTACTTGGGCAAGGGCGTTTTGAAGGCGAGTTGAAACGTC	360
DB	301	CTGCGTGAACGGTGGCGATGCTCTACTTGGGCAAGGGCGTTTTGAAGGCGAGTTGAAACGTC	360
QY	361	AACGAAGAANAATCCGGCGACGAGCTCGCTGGCCTTAGAGGCTGACGATCAGGGCTCATCGAC	420
DB	361	AACGAAGAANAATCCGGCGACGAGCTCGCTGGCCTTAGAGGCTGACGATCAGGGCTCATCGAC	420
QY	421	GAAGCAATGATCAACGCTTTGATGGCAACCGCAACAAAGTCCCGCTCGGTGCAACCAATC	480
DB	421	GAAGCAATGATCAACGCTTTGATGGCAACCGCAACAAAGTCCCGCTCGGTGCAACCAATC	480
QY	481	CTTGGTGTTCATGAGCTTTGTCAAAGGCTGCTGTGATTCGCGAGGCTCCCACTGTTC	540
DB	481	CTTGGTGTTCATGAGCTTTGTCAAAGGCTGCTGTGATTCGCGAGGCTCCCACTGTTC	540
QY	541	CGCTACATCGGTGGACCAAAACGACACGTTCTTCAGTTCCATGATGAACATCATCAAC	600
DB	541	CGCTACATCGGTGGACCAAAACGACACGTTCTTCAGTTCCATGATGAACATCATCAAC	600
QY	601	GGTGGCGCTCACGCTGACTCCGGGTGTGACGTTCAGGAATTCATGATCGCTCCCATCGGT	660
DB	601	GGTGGCGCTCACGCTGACTCCGGGTGTGACGTTCAGGAATTCATGATCGCTCCCATCGGT	660
QY	661	GCAGAGACTTTCTCTGAGGCTCTCCGCAACCGCGCGAGGCTCTACCGACACTGAAGTCC	720
DB	661	GCAGAGACTTTCTCTGAGGCTCTCCGCAACCGCGCGAGGCTCTACCGACACTGAAGTCC	720
QY	721	GTCCATCAAGGAAAAGGGGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTC	780
DB	721	GTCCATCAAGGAAAAGGGGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTC	780
QY	781	GTGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTC	840
DB	781	GTGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTC	840
QY	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTTCCTCAAGGAC	900
DB	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGAGGTTGCTTCTCTGAGTTCCTCAAGGAC	900
QY	901	GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGAGCTGAGATGGCAAAAGTTTACGCT	960
DB	901	GGCACCTACCACTTCGAAGGTGGCGAGCACTCCGAGCTGAGATGGCAAAAGTTTACGCT	960
QY	961	GAGCTCGTTGACGGGTACCCCAATCGTCTCCATCGAGGACCCACTTCGAGGAGATGACTGG	1020
DB	961	GAGCTCGTTGACGGGTACCCCAATCGTCTCCATCGAGGACCCACTTCGAGGAGATGACTGG	1020
QY	1021	GAGGTTTACACCAACCTCAACCGCAACCACTCGGCGACAGGTTCCAGATCGTTGGCGACGAC	1080
DB	1021	GAGGTTTACACCAACCTCAACCGCAACCACTCGGCGACAGGTTCCAGATCGTTGGCGACGAC	1080
QY	1081	TTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCAATCGTAAAGAGGCTGCCAATCC	1140
DB	1081	TTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCAATCGTAAAGAGGCTGCCAATCC	1140
QY	1141	ATCCTGGTTAAGGTGAACCAAGATCGGTACCCCTCAACCGAGACCTTCGAGCTGTGCAATG	1200
DB	1141	ATCCTGGTTAAGGTGAACCAAGATCGGTACCCCTCAACCGAGACCTTCGAGCTGTGCAATG	1200
QY	1201	GCTCAACCGCGAGGCTACACCTTCATGATGTCCACCGTTCGCGGTGAGACCGAGGACACC	1260
DB	1201	GCTCAACCGCGAGGCTACACCTTCATGATGTCCACCGTTCGCGGTGAGACCGAGGACACC	1260
QY	1261	ACCATTGCTGACTCGCAGTTGCACTCAACTGTGCGGCAAGTCAAGACTGGTGTCTCCAGCA	1320

Db	1261	ACCAATTGCTGACTCCGCAAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA	1320
QY	1321	CGTTCCGACCGTGTGCGAAAGTACAACACAGCTTCTCCGCATCGAGCAGCTGCTTTGGCGCAC	1380
Db	1321	CGTTCCGACCGTGTGCGAAAGTACAACACAGCTTCTCCGCATCGAGCAGCTGCTTTGGCGCAC	1380
QY	1381	GCCGGCGTCTACGCAAGTTCGAGCGCATTTCCACGCTTTTCAGGCTTAATAAAGCGCTT	1440
Db	1381	GCCGGCGTCTACGCAAGTTCGAGCGCATTTCCACGCTTTTCAGGCTTAATAAAGCGCTT	1440
QY	1441	TTTCAGCGCCGGTAAACCTCAAGGTTCGCGGGCGTGGTTCCTTACTACTGTTACTGTTGT	1500
Db	1441	TTTCAGCGCCGGTAAACCTCAAGGTTCGCGGGCGTGGTTCCTTACTACTGTTACTGTTGT	1500
QY	1501	GACTATGATCGAGGATTTATGCAAGACAGAAAGAAAATCATAAAGGCTTTGTTCTGTCT	1560
Db	1501	GACTATGATCGAGGATTTATGCAAGACAGAAAGAAAATCATAAAGGCTTTGTTCTGTCT	1560
QY	1561	CAAGCAGGGAACGTGCTT	1578
Db	1561	CAAGCAGGGAACGTGCTT	1578
RESULT 5			
LOCUS	AP005277	333150 bp	DNA
DEFINITION	Corynebacterium glutamicum ATCC 13032	linear	BCT 08-AUG-2002
ACCESSION	AP005277	BA000036	
VERSION	AP005277.1	GI:21323710	
KEYWORDS			
SOURCE	Corynebacterium glutamicum ATCC 13032		
ORGANISM	Corynebacterium glutamicum ATCC 13032		
REFERENCE	1		
AUTHORS	Nakagawa, S.		
TITLE	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 333150)	
AUTHORS	Nakagawa, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.		
	Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,		
	Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com,		
	Tel: 81-44-829-3031, Fax: 81-44-813-1651)		
	This sequence is conducted by collaboration of Kyowa Hakko Kogyo		
	Co. Ltd. And Kitasato University.		
COMMENT			

FEATURES

```

1. 353930
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
complement(63..1292)
/genes="Cgl0944"
complement(63..1292)
/genes="Cgl0944"
/codon_table=1
/transl_table=1
/product="Hypothetical membrane protein"
/protein_id="BAB98337.1"
/db_xref="GI:21323711"
/translation="MTNPTERNARLLIWANGLQNGDQIVAAKTVLPWLLQAAGAPG
FLALALVPIREAGSMPLPQAATIGVLRQTSRKVMVIGNSGQFVSALGIGVAAFLPRG
WALGITVYLLAALSVFVAGSSGIADKQVGKGLVTRATVIGVMVLGVL
ATAIPLGSHSPTRVLAAVFIACSSFSWLFASIVFARIYAKPATPKNAPSANPWRRCI
AALKDKAPRRFVLRSMNLVTALSTAFIVLAALAEESGNSIDSLGFLIASGLASWVG
RISQWSDHSKKNWAGGALFSGIVLILVLSAPAPQINTLVPFLSFLFLITLAHTA
IRVARTYVMWAEQDQTRRYADANTLIMGVVLLVIGALSQGFIAIFGNEALLFLAAI
GLGITTSARGLKEVSAG"
1382..2917
gene

```

```

/ gene="Cg10945"
/ db_xref="GI:21323712"
/ protein_id="BAB98338.1"
/ product="Putative multicopper oxidases"
/ translation="MPELNRTFFKGVLAATVVGAVQLVACSSDDVGVGGEPRTL
PIPADLGTREGSVHFALEAQGESQILPDVTTKTWGFNGTHLPTLVVKKGDVHV
DVNNIDEMTTVMHGMKLPALADGGPHSPIGQOTWSPMTWVANDAAATLWHPHTHG
LTGLHAYRGLAGMLIVDEATDKLDLPREYGVDDIPLVLMDBRPLEDGSIDEDLDPL
EPQDRTTAINGTNAHEDATRRVRFLVNGSNRFFYNLAFSDTRTFQVFIASGGLLD
GLGDTTFLATIGPERWEIVVEPEGEDVLESVGFEDNYGVDDDFVPDFGMSDSFOL
LITTPGDDAAQAPALPGVLVKSTEPDIVDATERFTIMNTFSINDLQMDQRVDVVVD
HDQPEVMVINDNSDMPNPHVHDARFKVLKFEQTDVELFNDGMDKDTVGLPPGATATL
AVREGHYPDPQMPFMYTHCHMYLHEDQGMQGFVIVERDEPAAVLGSGTSSIDSAGG
HAH"
/ gene="Cg10946"
/ db_xref="GI:21323713"
/ protein_id="BAB98339.1"
/ product="ABC-type transporter, ATPase component"
/ translation="MTTTPAIDVTDLVRTGYDYTAVGLNFHVORGVFGLGTNGAG
KTSLEVIKGLSAPSGTVRISGLDPVADRAILPELIGMLQSGLESQTLVARTWDM
WHGCTVPRRAIKDVLADVDLHRENVKVGALSGEQRRLDLACALLGDPSTLFDLEPT
TGLDPERRRHTMQLLDLKGQVMTMLTTHYLEABFLCDRIAMNAGEIAVEGTIDE
LVAREKSIISFVLGGQVELPVLGSAEIIIRDNNHVRITATTLQHTLEILTAAETGI
ALEGFAAKPATLESVFMIDIASLENTSLQTA"
/ gene="Cg10947"
/ db_xref="GI:21323714"
/ protein_id="BAB98340.1"
/ product="ABC-type transporter, permease components"
/ translation="MTTSHTARGIQAHPERKTSFPFKTSLFKAEMLQPRNKTLLEW
ATVPFVGLPFLIIGNGAAASNFDFVMTLLFVQFTVLSMATHRDERVLKR
LRTGEARDIDIGAIKCPGALLTFTVVIPLLVGLGAPAPINLVIFAVLIGLLL
CSALALMTSGFRNAEAQMTSMFVFLMAGLGSIRFPQDGSIVADILAYTPPAIS
DLVQIGWAGATFADSVGVGEAANFAGIQDMLIPLGLIAATAAANAANRYMRWDSY
R"
/ gene="Cg10948"
/ db_xref="GI:21323715"
/ protein_id="BAB98341.1"
/ product="Two-component system, sensory transduction
histidine kinases"
/ translation="MKNQFWTAGTWARWFSRGSVLLASPVTAPLNWRRLPNLAKYTL
YTRVSLQAIPIVLLSAYFLGIVANAGTNPSPFVLLGFSVILLIVLVVYEQFSLNS
HPRRSVQPPFFTCGLVNLVGVVSVVQLIPLGMSDNTRATALIFTLCVFLISYAI
PMWYRWVLIAMSAVLNWTSTTDDYLSALWVVPPLMAGTIVRLSVMTVDVMEKVERS
RELEASLRVTERLRPAQELHDTLGOHLAAMSVKSELALAKAGDDRLDENELELOK
LRTSMSEMRDVSGVRYTNVLALETEGAKSLADAHILHSVIGTTSQVSPAPRELCAW
LVREATTNLRHSDATDATTLSSTEVRMNNGVNKDIGRLSGLSAPSAESAGMTL
IVSREDDQFSVRMLINAPANTPAKEA"
/ gene="Cg10949"
/ db_xref="GI:21323716"
/ protein_id="BAB98342.1"
/ translation="MISIAADDEALIASLTLLEPDLVVRPTAGSGEELIETWA
DPSNRDVCVLQGLGIDGIDTATLMTTDLAVLIVTSHARPQLKALAGVLG
FLPRTADAFPAITRTVHAGRIYIDPELAAMTISAGESPLTNREEVLELACQGLSA
EIIAAHLAPGTRNYLSQAMTKVQAQRFAETFRARELGWL"
/ gene="Cg10950"
/ db_xref="GI:21323717"
/ translation="MKHYMSYFAGACPLNTPHTKYVFMWDIVSVNHFWSNTNNWN
MTGLPRETVPAFVPGSVTLAVLSF"
/ gene="Cg10951"
/ db_xref="GI:21323718"
/ translation="MTAWLCSPPSGVQDEYDMNGVVQPOEHLDATLIAADFHGNPN
SGDKRRLNFOQKRYALNRTVRDVPDGLDLAALLTFPSILSIAPAVLLGYVITIF
LADSTELNLNVRDENVQYVPEDQSHVGVNDISAGSAAAGQGVAVGVITALTSS
AVYRPFRCNAVYGRSGRTLLIKRWMLLFLNALLLGLIIILSVVNLNETLVMGIF
APYAEPLHNTNLSFLTDFMPFIWVRPFVIVGLIMFVATLYWAPNARPMKFRWL
SLGSLAI VIGILLAGVGLNFYFTLFAAFSSYGAVGSLLA VFIALWVFNICLIIGLKID
VEISRAKQLQAGMCAEDYSILVPSRSIEKVAKMQRQORLMDQAAARESN"
/ gene="Cg10952"
/ db_xref="GI:21323719"
/ protein_id="BAB98345.1"
/ translation="MWRTHASSILRMRORGVAFQAFGLGMQLSEALVPIAI
GLIVDHAVLTKDLRRLVGLVAVFLVFLVPSIRFGSRAALNRAVNFESRALRVEVAD
HALNLDPRNLVFEVNSRSTADASTRIFGQIGTVSAATFLGTAATYLLISDLV
GLLVLVLPITISGVVALAKSISKRSTVQOEKLAESGAQASDIMMGLRVTIKAIIGGERW
AVTTFKASQASRAAAVDTAVASGVAGIGELSIANVLAALLAGNRVTTGELGPGQ
LIAIVGAVVLSPIRLNSINASIAHGAERVANFLNLDSEQAQVYESSETINDGE
FLVIVPASTLPHGDNILATPHADPEGLTSLRSNHNHEDNVPIDPQVIRASULTDI

```


AUTHORS Kalinowski, J.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer Genomforschung, Universitaet Bielefeld, Universitaetsstrasse 25, 33615 Bielefeld, Germany
COMMENT E-mail: Joern.Kalinowski@Cebitec.Uni-Bielefeld.DE
 This sequence was accomplished by collaboration between Degussa AG and Bielefeld University.
 join(BX927148.1.1..348071, BX927149.1.1:51..349887, BX927150.1.1:51..348475, BX927152.1.1:51..349799, BX927153.1.1:51..349584, BX927151.1.1:51..349459, BX927155.1.1:51..349736, BX927156.1.1:51..349115, BX927157.1.1:51..140057).
FEATURES Location/Qualifiers
 1..348475
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="Genomic DNA"
 /strain="DSM 20300 = ATCC 13032"
 /db_xref="taxon:196627"
 /note="IS fingerprint type: 4-5"
 complement(381..1610)
 /locus_tag="cg0783"
 complement(381..1610)
 /locus_tag="cg0783"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="CAF19386.1"
 /db_xref="GI:41324905"
 /translation="MAEKGIIVPKLSLTGEGDYFTLWPNWREGNEWQAFGLGANDDLY VFNPSPELLVEESDDHDLTSPHKFNFGSDAARVVPDANAEIDLIIGAPALLADRP SYANVKTLRIFAVMSLGHVTAASPVTFVSSHVSLSNVDRGSEHSGPGLGEMTS VGRAVATNNGVADLBALITVKEVNSLITFEDAAITRIQEAQAAVQAAAEAEQIAKEE AEKVDYDNSMAAGIDPIKVISDGRITVTLRTYLGQPVFKGFGIEITPFSOKSL LRNLVHDDHDLAASWSDMLGINSGETELLVHSNVSFNGLVKIDINTSDVAVDT QOMARVHDLADATWDDMADDSNNYSIFLANFRMQDIYSYMLGGSSTSGYVHPAPNDHS ESWRELEMLIKRFSKF"
 complement(1677..2537)
 /locus_tag="cg0784"
 complement(1677..2537)
 /locus_tag="cg0784"
 /codon_start=1
 /transl_table=11
 /product="PUTATIVE CELL WALL-ASSOCIATED HYDROLASE"
 /protein_id="CAF19387.1"
 /db_xref="GI:41324906"
 /translation="WIGLISLSRLATHEPAQLPOLAIPOAPDISAALGRELISNP TLLSLFATSQQHLLVTSLLIQAPLIDARHDIATTAOHLVNOAGLVMSFSMPN AESFAARTELISLPLMTAEASQRUEMSRQVPIVEKLVQIQILTEPTAEPAVETA PVYMAQAPSDNEVGOAAVAKEALGTPLYMGTSVSGPDCSLGTOMAWRRAGVEI PRIADQAVGRQVYDELQEGDLTIWDGHVAMYAGGQIIEAGDPVQLNPRTSNIG AFHGYFRPTG"
 complement(2534..2797)
 /locus_tag="cg0785"
 complement(2534..2797)
 /locus_tag="cg0785"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein predicted by Glimmer/Critica"
 /protein_id="CAF19388.1"
 /db_xref="GI:41324907"
 /translation="NRINLPHAKELAHLCILPTPAVPALPTDGAQPDIIHQALSASI ATYARNLLLSHTAENLGNRLTGLAEIETDDQLAHALERLT"
 2896..3531
 /gene="upp"
 /locus_tag="cg0786"
 2896..3531
 /gene="upp"
 /locus_tag="cg0786"
 /EC_number="2.4.2.9"
 /codon_start=1
 /transl_table=11

gene
CDS
 /product="PUTATIVE URACIL PHOSPHORIBOSYLTRANSFERASE"
 /protein_id="CAF19389.1"
 /db_xref="GI:41324908"
 /translation="MDITIVNHLVVASRLTLRLDRSDNAAAFRAANDLGAMLIYEAS RDLVEHPDTTPVAMAGTRLKOPPIIVPIIRAGLMDIPALSMIPDAQVGFIGLAR DEETHEPVYLEALPQDLSNQPVFLVDPMLATGGSLHAIIRLLADRGATDITACMVS AQPGVDALESGLPVLVLTATIDPGLDENAVIVPGLGDAGRLYGRNIDL"
 3625..3999
 /locus_tag="cg0787"
 3625..3999
 /locus_tag="cg0787"
 /codon_start=1
 /transl_table=11
 /product="transcriptional regulator"
 /protein_id="CAF19390.1"
 /db_xref="GI:41324909"
 /translation="MSEKRWSSYGHAFASRLKRLTLRGFSOEELADLSGVSRRNTIS VERNENKGNADVDPOLSNIRYLAQALDVPPIALMPAGSVPAKICVDETAADIVRWPS EKDPILLFDADLALTRRQPGN"
 complement(3996..5588)
 /gene="pmmB"
 /locus_tag="cg0788"
 complement(3996..5588)
 /gene="pmmB"
 /locus_tag="cg0788"
 /EC_number="5.4.2.2"
 /EC_number="5.4.2.8"
 /codon_start=1
 /transl_table=11
 /product="PHOSPHOGLUCOMUTASE/PHOSPHOMANNOMUTASE"
 /protein_id="CAF19391.1"
 /db_xref="GI:41324910"
 /translation="MDESQLSFGTAGLRAPVGPAPKQNVQLQVTRTTAGVASWLAER AALNPVPHVPEDETGIGALYPOQGLVGVYDARYGSHFTATTAETVAGAGFEV TLLPTSPLEPLIPWLVNKGDLGQVITASHNGAADNGYKFLSNGRGLYSEGLEPEL AHNAVEDPIRVPRVTRPTADQLRKYVDEMVSIVTPDQADLLRVNBSRGNLRVVTTA LHVGGRAMANAQFAPAGPPTHGVKAQOYPDPTFTVAPNPEEPGSAIELLERAKEK NADILPALDPADRCAGVIRTAGGHRMLSGDEVTLLATRLVPEYSGRGPVAVATT VVSSQLIGIIADKGDYSETLTGPKNLSRAADGLDGLAFAYEAEVAGTCPPVDVVD KGGISTALPMASWAAELKQASLOOKNELRYRYGYFPASSQAVTSTSPRELVDHWI AHPQELIGVSVTHILPEKQIALHGQVGHVHIRAIGRVSGTEAKAKLYLEVQQASS HUEAAQLHOLEDEVOQSLSKL"
 5693..6880
 /gene="amiA"
 /locus_tag="cg0789"
 5693..6880
 /gene="amiA"
 /locus_tag="cg0789"
 /codon_start=1
 /transl_table=11
 /product="PUTATIVE N-ACYL-L-AMINO ACID AMIDOHYDROLASE"
 /protein_id="CAF19392.1"
 /db_xref="GI:41324911"
 /translation="NMEIGVQVAMNDRHHEDEVIKWRHHLHSHPELSHMEYRTEYLA SVLKDHMBPHLPFGTGLMVDIGEGDSRLAFRADIDALPLETESTGLEFSGTATGVAH ACCHDYVTALALACALANTIELPIGIRVIFQPAEVMTCGATDVAHGLDGDVDAIY AIHVPEKLVGRVGVZAGALTSASDVIEIRVKGEGGSAPHSADVVYALSILVVDL PGLSRKVDPTGTVLVFGTINAGYAPNALPDSGIVSGTIRTDISTWRDMLISEL VEQVLTPTGVTHELIVPGVPVINDVDVATALLASARDMDTSVVQVAPQSGGSEDFS WYLEHVPGSMARLGCWPGHKQDLHQSDLVVDERAIGVGRVLGSLVQOQYSRSEAF LNS"
 7020..8429
 /gene="lpda"
 /locus_tag="cg0790"
 7020..8429
 /gene="lpda"
 /locus_tag="cg0790"
 /EC_number="1.8.1.4"
 /codon_start=1
 /transl_table=11
 /product="DIHYDROLIPOAMIDE DEHYDROGENASE"
 /protein_id="CAF19393.1"
 /db_xref="GI:41324912"

```

/translation="MAKRIVIIGGPAGYEALALAGAKYGAETVIVIEDVGVGSAVTMD
CVPKSFIAIGTIGKIDLRADDMGLNRGLGKSHLEIDALNIRVKOLAKAQSEDILOL
QSDSVNMGNGFDDYNTKQOTTHYIKVTHSDGSEBETVEDILVLATGATPRILKGA
PDGERLTWROVYDIEELPHLIVGSGVTGAEFVAFELGVKVTMVAARDLILPHD
DADAADVLEAERSVLEKHARVESVTRTEGGVVCVTRADGREIYGHSHALMTVGS
PNTADLGNIGVELAPSCHKVDRVSRNIPGVYAAGDCTDLFFPLASVAAMQGRIAM
YHALGVSFPIRLKTVATVAFTRPEIAAAGVTHAQVDSGEVSARVILVPLATNPRAKM
RSLRHGFKLFCRNSGLIIGVVVAPTASELILPITAVAVTNRLTAVDLADTFAVYPS
LGSITEAARQVQHQDLG"
8826..12248
/gene="pyc"
/locus_tag="cg0791"
8826..12248
/gene="pyc"
/locus_tag="cg0791"
/EC_number="6.4.1.1"
/codon_start=1
/transl_table=11
/product="PYRUVATE CARBOXYLASE"
/protein_id="CAP19394.1"
/db_xref="GI:41324913"
/translation="MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPR
EDRGFHRSPASAVRIGTEGSPVKAYLDIDELIGAAKKVADADAIIPGYFLSENAQL
ARECAENGITFGPTDEVLDLTDKSRVATAKAGLPLVAESTPSKNIDEIVKABG
QTPYIFKAVAGGGGRMRFVAPDELRLKLAATEASREAEAFDGDVAVYVERAVINPQH
IEVQILDHTGTVHLYERDLSLORRHKVIEAPAOHLDPRLDRICADAVKFKRSI
GYGAGTVEFLVDEKGNHVPFIEPNRIQVHTVTEVTDLVKAOMRLAAGATLKEI
GLTDKIKTHGAALQCRITTEDPNNGFRPDGTGIIYAIRSPGGAGVLDGNAQLGGEIT
AHFDSMLKMTCRGSDFFETAVARAQRALAEFTVSGVATNIGFLRLALBEDFTSKRIA
TGFIDHPHLAPPDARQGRILDYLDADVTNKPGRVPRKDVAAAPIDKLPNIKOLPL
Query Match
Best Local Similarity 99.9%; Score 1576.4; DB 1; Length 348475;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 GGCTGGGATATGGTGTAGTTTCGCCCACTAAATTTCAACTGATTCCTCATCGAAAACAAGA 60
Db 338411 GGCTGGGATATGGTGTAGTTTCGCCCACTAAATTTCAACTGATTCCTCATCGAAAACAAGA 338470
61 TTCCTGCAACAATTTGGGTGTAGACGTGATTAAGACATTTTGAATCAGTGAATTAATTCATG 120
Db 338471 TTCCTGCAACAATTTGGGTGTAGACGTGATTAAGACATTTTGAATCAGTGAATTAATTCATG 338530
121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTCGTCTCG 180
Db 338531 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTCGTCTCG 338590
181 GAAATTTCTCGATCTCCCGCGGTAAACCAACCGTTCGAGGCGAGAGGTTTTCCTGGATGACGT 240
Db 338591 GAAATTTCTCGATCTCCCGCGGTAAACCAACCGTTCGAGGCGAGAGGTTTTCCTGGATGACGT 338650
241 TCCACGGTGTCCGAGGTGTCCATCCGGCGCATCCACCGCGGTCCACGAGGCTCATGAG 300
Db 338651 TCCACGGTGTCCGAGGTGTCCATCCGGCGCATCCACCGCGGTCCACGAGGCTCATGAG 338710
301 CTGCGTGAAGTGGCGCATCTGCTACCTGGGCAAGGGCGTTTGAAGCAGTTGAAACAGTC 360
Db 338711 CTGCGTGAAGTGGCGCATCTGCTACCTGGGCAAGGGCGTTTGAAGCAGTTGAAACAGTC 338770
361 AACGAGAAATCGGCGACAGCTCGTGGCCCTAGAGCTGACGATCAGCGCTCATCGAC 420
Db 338771 AACGAGAAATCGGCGACAGCTCGTGGCCCTAGAGCTGACGATCAGCGCTCATCGAC 338830
421 GAAGCAATGATCAAGCTTGTATGGCAACCGCAACAGTCCCGCTGGGTGCAACACGCAATC 480
Db 338831 GAAGCAATGATCAAGCTTGTATGGCAACCGCAACAGTCCCGCTGGGTGCAACACGCAATC 338890
481 CTTGTGTTTCCATGCTGTTCGAAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTTC 540
Db 338891 CTTGTGTTTCCATGCTGTTCGAAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTTC 338950
541 CGCTACATCGGTGGACCAACGCAAGCTTCTCCAGTTCCATGATGATGATCAATCATCAAC 600

```

```

Db 338951 CGCTACATCGGTGGACAAACGACACAGCTTCTTCCAGTTCCAATGATGAACATCATCAAC 339010
Qy 601 GGTGGCGCTCACGCTGACCTCGGTGTGACGTTGACGTTTCAAGAAATTCATGATCGTCCAAATCGGT 660
Db 339011 GGTGGCGCTCACGCTGACCTCGGTGTGACGTTTCAAGAAATTCATGATCGTCCAAATCGGT 339070
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCCAACGGCGCGGAGGTCTACCAACGCACTGAAGTCC 720
Db 339071 GCAGAGACCTTCTCTGAGGCTCTCCGCCAACGGCGCGGAGGTCTACCAACGCACTGAAGTCC 339130
Qy 721 GTCATCAAGGAAAAGGCGCTGTCCACCGGATTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Db 339131 GTCATCAAGGAAAAGGCGCTGTCCACCGGATTTGGCGATGAGGCGGCTTCGCTCCTTCC 339190
Qy 781 GTCGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTTAAAGCAATCGAAGGCTGCGCTTC 840
Db 339191 GTCGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTTAAAGCAATCGAAGGCTGCGCTTC 339250
Qy 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGAGCTTGTCTCTGAGTTCTTCAAGGAC 900
Db 339251 ACCCAGGCAAGGACATCGCTCTTGCTCTGGAGCTTGTCTCTGAGTTCTTCAAGGAC 339310
Qy 901 GGCACCTACCACTTCCAAAGGTGGCCAGCACTCCGAGCTGAGATGCGAAACGTTTACGCT 960
Db 339311 GGCACCTACCACTTCCAAAGGTGGCCAGCACTCCGAGCTGAGATGCGAAACGTTTACGCT 339370
Qy 961 GAGCTCGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
Db 339371 GAGCTCGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG 339430
Qy 1021 GAGGCTTACCAACCTCACCGCAACCACTCGGCGCAAGGTTTCAGATCGTTGGCGACGAC 1080
Db 339431 GAGGCTTACCAACCTCACCGCAACCACTCGGCGCAAGGTTTCAGATCGTTGGCGACGAC 339490
Qy 1081 TTCTTCTGTCACCAACCTCGAGCGCTGAGAGGCGCATCGTAAAGAGGCTGCAACTCC 1140
Db 339491 TTCTTCTGTCACCAACCTCGAGCGCTGAGAGGCGCATCGTAAAGAGGCTGCAACTCC 339550
Qy 1141 ATCTTGGTTAAGGTGAACCGAGATCGGTACCTTCAACCGAGACCTTCGACGCTTCGACATG 1200
Db 339551 ATCTTGGTTAAGGTGAACCGAGATCGGTACCTTCAACCGAGACCTTCGACGCTTCGACATG 339610
Qy 1201 GCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTCCGCTGAGACCGAGGACACC 1260
Db 339611 GCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTTCGCTGAGACCGAGGACACC 339670
Qy 1261 ACCATTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA 1320
Db 339671 ACCATTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA 339730
Qy 1321 CGTTCGACCGGTTCGCAAAAGTACAAACGAGCTTCTCCGATCGAGCAGCTGTTGGCGAC 1380
Db 339731 CGTTCGACCGGTTCGCAAAAGTACAAACGAGCTTCTCCGATCGAGCAGCTGTTGGCGAC 339790
Qy 1381 GCGCGGCTTACGAGGCTGCGAGCGCATTCACGCTTTCAGGGCTAAATAAAGCGCTT 1440
Db 339791 GCGCGGCTTACGAGGCTGCGAGCGCATTCACGCTTTCAGGGCTAAATAAAGCGCTT 339850
Qy 1441 TTGACGCGCGGTAACTCAAGGTTGCGGCGCTGTTGCCCTTACTACTGTTACTGTTGT 1500
Db 339851 TTGACGCGCGGTAACTCAAGGTTGCGGCGCTGTTGCCCTTACTACTGTTACTGTTGT 339910
Qy 1501 GACTATGATCGAGGATTAAGCAAGGACGAGAAAGAACTCATAAAGGCGCTTGTTCCTGCTCT 1560
Db 339911 GACTATGATCGAGGATTAAGCAAGGACGAGAAAGAACTCATAAAGGCGCTTGTTCCTGCTCT 339970
Qy 1561 CAAGCAGGGAACGTGCTT 1578
Db 339971 CAAGCAGGGAACGTGCTT 339988

```

RESULT 7
AX1217146

LOCUS AX127146 349980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 7062 from Patent EP1108790.
ACCESSION AX127146 AX114121
VERSION AX127146.1 GI:14041134
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE 1
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K., Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7062 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source Location/Qualifiers
1..349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/notes="Seq 1 to long (3.309.400) split in 11, seq 7062
0.900.001 1.249.980"
ORIGIN
Query Match 99.9%; Score 1576.4; DB 6; Length 349980;
Best Local Similarity 99.9%; Pred. No. 4.8e-289;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCTGGGGATATGGTATGTTTCGCCACTAATTTCAACTGATTCGCTCATCGAACAAGA 60
DB 134799 GGCTGGGGATATGGTATGTTTCGCCACTAATTTCAACTGATTCGCTCATCGAACAAGA 134858
QY 61 TTGCTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATCAGCTGAATATTTAG 120
DB 134859 TTGCTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATCAGCTGAATATTTAG 134918
QY 121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATATCAATGCAATTCGCTCGC 180
DB 134919 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATATCAATGCAATTCGCTCGC 134978
QY 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGAGGT 240
DB 134979 GAAATTTCTGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGAGGT 135038
QY 241 TCCACCGGTTCGACAGGTGTTCAATCCGGCGCATCCACCGCGCTCCACAGGCTCATGAG 300
DB 135039 TCCACCGGTTCGACAGGTGTTCAATCCGGCGCATCCACCGCGCTCCACAGGCTCATGAG 135098
QY 301 CTGCTGACCGGTGGCGATCGTACCTGGGCAAGGGCGTTTGAAGGCGATTTGAACACGTC 360
DB 135099 CTGCTGACCGGTGGCGATCGTACCTGGGCAAGGGCGTTTGAAGGCGATTTGAACACGTC 135158
QY 361 AAGCAAGAAATCGGCGACGAGCTGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420
DB 135159 AAGCAAGAAATCGGCGACGAGCTGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 135218
QY 421 GAAGCAATGATCAAGCTTGTATGGCAGCCGCAACAAAGTCCCGCTGGGTGCAACGCAATC 480
DB 135219 GAAGCAATGATCAAGCTTGTATGGCAGCCGCAACAAAGTCCCGCTGGGTGCAACGCAATC 135278
QY 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGCGCTCCCACTGTTTC 540
DB 135279 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGCGCTCCCACTGTTTC 135338
QY 541 CGCTACATCGGTGGAACCAACGACACGCTTCCTCAGTTCCCAATGATGAACATCATCAAC 600
DB 135339 CGCTACATCGGTGGAACCAACGACACGCTTCCTCAGTTCCCAATGATGAACATCATCAAC 135398
QY 601 GGTGGGCGCTCAGCTGAGCTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT 660
DB 135399 GGTGGGCGCTCAGCTGAGCTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT 135458
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGCACTGAAGTCC 720

DB 135459 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACCAACGCACTGAAGTCC 135518
QY 721 GTCATCAAGGAAAGGGCTCTGTCCAGGACTTTGGCGATTTGGCGGCTTCGCTCCTTCC 780
DB 135519 GTCATCAAGGAAAGGGCTCTGTCCAGGACTTTGGCGATTTGGCGGCTTCGCTCCTTCC 135578
QY 781 GTGGCTTCCACCCGTCGAGGCTCTTGAACCTTATCGTTAAGGCAATCAGAAAGGCTTTC 840
DB 135579 GTGGCTTCCACCCGTCGAGGCTCTTGAACCTTATCGTTAAGGCAATCAGAAAGGCTTTC 135638
QY 841 ACCCGAGGCAAGGACATCGCTCTTGTCTGAGGCTTGTCTCTCTGAGTCTTTCAGAGGAC 900
DB 135639 ACCCGAGGCAAGGACATCGCTCTTGTCTGAGGCTTGTCTCTCTGAGTCTTTCAGAGGAC 135698
QY 901 GGCACCTTACCACCTTCAAGGTGGCGAGCACTCCGAGCTGAGATGCAAAACGTTTACGCT 960
DB 135699 GGCACCTTACCACCTTCAAGGTGGCGAGCACTCCGAGCTGAGATGCAAAACGTTTACGCT 135758
QY 961 GAGCTCGTTGAGCGCTTACCACCTTCCCAATTCGCTTCCATCGAGGACCACTGCGAGGAATGACTGG 1020
DB 135759 GAGCTCGTTGAGCGCTTACCACCTTCCCAATTCGCTTCCATCGAGGACCACTGCGAGGAATGACTGG 135818
QY 1021 GAGGCTTACACCAACCTTACCGCAACCATCGGCGACCAAGTTTCAGATCGTTGGCGAGGAC 1080
DB 135819 GAGGCTTACACCAACCTTACCGCAACCATCGGCGACCAAGTTTCAGATCGTTGGCGAGGAC 135878
QY 1081 TTCTTCTGTCACCAACCTTGAAGGCGCTTGAAGAGGCGATCGTGAAGAGGCTGCAACTCC 1140
DB 135879 TTCTTCTGTCACCAACCTTGAAGGCGCTTGAAGAGGCGATCGTGAAGAGGCTGCAACTCC 135938
QY 1141 ATCTCTGTTAAGGTGAACAGATCGGTACCTTCCACCGAGACCTTTCGACGCTGTCGACATG 1200
DB 135939 ATCTCTGTTAAGGTGAACAGATCGGTACCTTCCACCGAGACCTTTCGACGCTGTCGACATG 135998
QY 1201 GCTCACCGGCGAGGCTTACACCTTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 1260
DB 135999 GCTCACCGGCGAGGCTTACACCTTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 136058
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTTGTGCTCCAGCA 1320
DB 136059 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTTGTGCTCCAGCA 136118
QY 1321 CGTTCGACCGGTTCGCAAAAGTACAAACGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
DB 136119 CGTTCGACCGGTTCGCAAAAGTACAAACGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 136178
QY 1381 GCGCGGTCTACGCGAGGTGCGAGGCAATTCACCGCTTTCAGGGCTAAATATAAGGCGCTT 1440
DB 136179 GCGCGGTCTACGCGAGGTGCGAGGCAATTCACCGCTTTCAGGGCTAAATATAAGGCGCTT 136238
QY 1441 TTGACGCGCGGTAAACCTCAAGTTTCGCGGGCGTGTGCTTACTTACTGTTACTGTTGT 1500
DB 136239 TTGACGCGCGGTAAACCTCAAGTTTCGCGGGCGTGTGCTTACTTACTGTTACTGTTGT 136298
QY 1501 GACTATGATCGAGGATTTATGGCAAAAGCAGAAAGAAAACTCATAAAGGCTTGTTCCTGCT 1560
DB 136299 GACTATGATCGAGGATTTATGGCAAAAGCAGAAAGAAAACTCATAAAGGCTTGTTCCTGCT 136358
QY 1561 CAAGCAGGGAACGTCGCTT 1578
DB 136359 CAAGCAGGGAACGTCGCTT 136376

RESULT 8

AX763190
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX763190
Sequence 7 from Patent WO03040291.
AX763190
AX763190.1 GI:32257758
Corynebacterium glutamicum
Corynebacterium glutamicum

1405 bp
DNA
linear
PAT 25-JUN-2003

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1
Zelder, O., Pompejus, M., Schroeder, H., Kroeger, B., Kloppe, C. and
Haberhauer, G.
Genes encoding for carbon metabolism and energy-producing proteins
Patent: WO 03040291-A 7 15-MAY-2003;
BASP AKTINSELSCHAFT (DE)

FEATURES

source

1..1405
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..1378
/note="unnamed protein product; RXA00235"
/codon_start=1
/translation="MAEIMHVPAREILDRSGNPTVEAEVFLDQSHGVAGVPSGASTG
VHEAHLDDGGDRYLGKLVKAVNNVBEIGDELADLORLIDRAMIKLDGTANK
SRLGANALIGVSMVAKAADSAGLPLFRYIGGNHVLVPMNMLITGGHADSVD
VQEFMIAPGAEFSEALRNAGVHALKSVIKELSTGLGDEGGPAPSVGSTRAL
DLIVAIKAGTTPGKIALALDVASSFFKDDTHFEQGGHSAEAMVYAEILVDAY
PIVSTEDLQEDDMWGYTNLTATIGDKQIVGDDFFVFNPERLKEGIAKAAISILVK
VNQIGLTETFDVADHARAGYTSNMHSRSGEETDITADLAVALNCQIKTGPARS
DRVAKYNQLLRILEQLDGVVYAGRSAPRFG"

CDS

ORIGIN

Query Match 88.8%; Score 1401.8; DB 6; Length 1405;
Best Local Similarity 99.9%; Mismatches 0; Indels 0; Gaps 0;
Matches 1403; Conservative 0; Matched 2; Indels 0; Gaps 0;
51 CGAAACAAGATTCGTCGAACAATTGGGTAGACGCTGATTGAAGACATTTGATCAGTGA 110
1 CGAAACAAGATTCGTCGAACAATTGGGTAGACGCTGATTGAAGACATTTGATCAGTGA 60
111 ATAATCTAGTTAGTCCCAAGTTGGCATAGAGGCGCACAGTGGCTGAAATCATCAGCT 170
61 ATAATCTAGTTAGTCCCAAGTTGGCATAGAGGCGCACAGTGGCTGAAATCATCAGCT 120
171 ATTCGCTCGGAAATTCGCACTCCCGGGTAAACCAACCGTCGAGGAGAGGTTTCT 230
121 ATTCGCTCGGAAATTCGCACTCCCGGGTAAACCAACCGTCGAGGAGAGGTTTCT 180
231 GATGACGGTCCCAAGTGTGCGAGTGTTCATCCCGGCGATCCACGGGTCACGA 290
181 GATGACGGTCCCAAGTGTGCGAGTGTTCATCCCGGCGATCCACGGGTCACGA 240
291 GGCTCATGAGTGTGCGAGTGTGCGAGTGTTCATCCCGGCGATCCACGGGTCAGT 350
241 GGCTCATGAGTGTGCGAGTGTGCGAGTGTTCATCCCGGCGATCCACGGGTCAGT 300
351 TGAACACGTCAACGAAGAAATCGGCGACGAGTGTGCGGCTAGAGGTCAGCATCAGC 410
301 TGAACACGTCAACGAAGAAATCGGCGACGAGTGTGCGGCTAGAGGTCAGCATCAGC 360
411 CCTCATGACGAAGCAATGATCAAGTGTGATGGCACCGGCAACAAAGTCCCGGTCGGTGC 470
361 CCTCATGACGAAGCAATGATCAAGTGTGATGGCACCGGCAACAAAGTCCCGGTCGGTGC 420
471 AAACGCAATCTTGGTGTTCATGGCTGTTCGAAGGTCGTGATTCGGCAGGCT 530
421 AAACGCAATCTTGGTGTTCATGGCTGTTCGAAGGTCGTGATTCGGCAGGCT 480
531 CCCACTGTTCGCTACATCGGTGGACCAACCAACAGTGTTCGAGTTCGAATGATGA 590
481 CCCACTGTTCGCTACATCGGTGGACCAACCAACAGTGTTCGAGTTCGAATGATGA 540
591 CATCATCAACGGTGGGCTCAGCTGACTCCGGTGTTCAGCGTTTCAAGGATTCATGATCGC 650
541 CATCATCAACGGTGGGCTCAGCTGACTCCGGTGTTCAGCGTTTCAAGGATTCATGATCGC 600

QY 651 TCCAAATCGGTGACGAGACCTTCTCTGAGGCTCTCGCAACCGCGGAGGTTTACCAAGC 710
DB 601 TCCAAATCGGTGACGAGACCTTCTCTGAGGCTCTCGCAACCGCGGAGGTTTACCAAGC 660
QY 711 ACTGAAGTCCGTCATCAAGGAAAAGGCTCTGCAACCGGACTTGGCGATGAGGGCGGCTT 770
DB 661 ACTGAAGTCCGTCATCAAGGAAAAGGCTCTGCAACCGGACTTGGCGATGAGGGCGGCTT 720
QY 771 CGCTCCTTTCGGTCCGCTCCACCGCTGAGGCTTCTGACCTTATCGTTAAGGCAATCGAGA 830
DB 721 CGCTCCTTTCGGTCCGCTCCACCGCTGAGGCTTCTGACCTTATCGTTAAGGCAATCGAGA 780
QY 831 GGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTCTGCTCTGAGGCTTCTCTCTGAGTT 890
DB 781 GGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTCTGCTCTGAGGCTTCTCTCTGAGTT 840
QY 891 CTTCAAGGACGGGACCTTACCACTTGAAGGTGGGCGACACTCCCGAGCTGAGATGGCAAA 950
DB 841 CTTCAAGGACGGGACCTTACCACTTGAAGGTGGGCGACACTCCCGAGCTGAGATGGCAAA 900
QY 951 GCTTTACGCTGAGCTCGTTGACGGTACCCCAATCGTCTTCAATCGAGGACCCACTGCAGCA 1010
DB 901 GCTTTACGCTGAGCTCGTTGACGGTACCCCAATCGTCTTCAATCGAGGACCCACTGCAGCA 960
QY 1011 AGATGACTGGGAGGTTTACCAACCTCACCGCAACCATCGCGCAAGGTTTCAAGTCT 1070
DB 961 AGATGACTGGGAGGTTTACCAACCTCACCGCAACCATCGCGCAAGGTTTCAAGTCT 1020
QY 1071 TGGCGACGACTTCTGTCACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAAAGC 1130
DB 1021 TGGCGACGACTTCTGTCACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAAAGC 1080
QY 1131 TGCCAACTCCATCTCTGTTAAGGTGAACAGTCCGCTACCGAGACCTTCGAGC 1190
DB 1081 TGCCAACTCCATCTCTGTTAAGGTGAACAGTCCGCTACCGAGACCTTCGAGC 1140
QY 1191 TGTGACATGCTCACCGCGAGGCTACACCTCATGATGTCACCGCTTCGGGTGAGC 1250
DB 1141 TGTGACATGCTCACCGCGAGGCTACACCTCATGATGTCACCGCTTCGGGTGAGC 1200
QY 1251 CGAGGACACCAACCTGCTGACCTGCGAGTGTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
DB 1201 CGAGGACACCAACCTGCTGACCTGCGAGTGTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
QY 1311 TGCTCCGACGAGTTCCGACCGCTGCGAAGTACAAACAGGTTCTCCGATCGAGCAGCT 1370
DB 1261 TGCTCCGACGAGTTCCGACCGCTGCGAAGTACAAACAGGTTCTCCGATCGAGCAGCT 1320
QY 1371 GCTTGGCGAGCGCGGCTTACGAGGCTGCGAGGCGCATTCACCGCTTTCAGGGCTAAAT 1430
DB 1321 GCTTGGCGAGCGCGGCTTACGAGGCTGCGAGGCGCATTCACCGCTTTCAGGGCTAAAT 1380
QY 1431 AAAAGCGCTTTTTCGACCCCGGTAA 1455
DB 1381 AAAAGCGCTTTTTCGACCCCGGTAA 1405

RESULT 9
AX064945

LOCUS AX064945 1398 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 71 from Patent WO0100844.

ACCESSION AX064945

VERSION AX064945.1 GI:12542657

KEYWORDS

SOURCE Corynebacterium glutamicum

ORGANISM Corynebacterium glutamicum

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1

Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.

corynebacterium glutamicum genes encoding proteins involved in

carbon metabolism and energy production

Patent: WO 0100844-A 71 04 JAN-2001

JOURNAL

BASF AKTIENGESellschaft (DE)
 Location/Qualifiers
 1..1398
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101..1378
 /note="unnamed protein product; RXA00235"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC25712.1"
 /db_xref="GI:12542658"
 /translation="MAEIMHVPAREILDSRNPVEAEVFLDDSGHGVAGVPSGASTG
 VHEAHELDDGGRIYLGKVLKAVENNEEIGDELGAADQRLIDEMIKLDGTANK
 SLEGNATILGVSMAVAKAADSAGLPLFRYTGPNNAHLVPPMNNITGGHADSQVD
 VQEFMIAPIGAEITSEALRNGAEVTHALSVIKGSLSTGLGDEGGFAPSVSSTREAL
 DLIVAEIKAGPTPKDIALALDVASBPFDGTVHPFGGSHSAEMANVYARLVADY
 PIVSTEDPDDDEGVNLTATIGDKVQIVGDDPFVTPNPERLEKGIKAKAANSILYK
 VNQGTLETEDAVDMAHRAGYTSNMSHRSGETDITADIADVALNCGQIKTGAPARS
 DRVAKYNOLLRIEQLLDAGVYGRSAPFRPQG"

ORIGIN

Query Match 88.4%; Score 1394.8; DB 6; Length 1398;
 Best Local Similarity 99.9%; Pred. No. 1.9e-254;
 Matches 1396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 CGAAACAGATTCGTGCAACAAATGGGTGTAGACGTGATTAAGACATTTGATCAGGTGA 110
 DB 1 CGAAACAGATTCGTGCAACAAATGGGTGTAGACGTGATTAAGACATTTGATCAGGTGA 60

QY 111 ATATTCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCAGGT 170
 DB 61 ATAATTCCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCAGGT 120

QY 171 ATTGCTCGCAAAATCTCGACTCCCGCGGTAAACCCACCGTCCGAGGCGAGAGTTTCT 230
 DB 121 ATTGCTCGCAAAATCTCGACTCCCGCGGTAAACCCACCGTCCGAGGCGAGAGTTTCT 180

QY 231 GGATGACGGTTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCAGGT 290
 DB 181 GGATGACGGTTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCAGGT 240

QY 291 GGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 350
 DB 241 GGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300

QY 351 TGAATAACGTCAGCAAGAAATCGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCG 410
 DB 301 TGAATAACGTCAGCAAGAAATCGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCG 360

QY 411 CCTCATGCAAGAAAGATGATCAAGCTTGTAGTGGCAACCGCCAAAGTCCCGCTGGGTGC 470
 DB 361 CCTCATGCAAGAAAGATGATCAAGCTTGTAGTGGCAACCGCCAAAGTCCCGCTGGGTGC 420

QY 471 AAAGCAATCTTGGTGTTCATGGCTGTGTCGAAAGGCTGCTGCTGATTCGAGGCT 530
 DB 421 AAAGCAATCTTGGTGTTCATGGCTGTGTCGAAAGGCTGCTGCTGATTCGAGGCT 480

QY 531 CCCACTGTTCCGCTACATCGGTGGACCAAGCGCACAGTCTTCCAGTTCCTCAATGATGAA 590
 DB 481 CCCACTGTTCCGCTACATCGGTGGACCAAGCGCACAGTCTTCCAGTTCCTCAATGATGAA 540

QY 591 CATCATCAAGGTTGGCTGCTACGCTGATTCGGTGTGAGCTTCCAGGAATTCATGATCGC 650
 DB 541 CATCATCAAGGTTGGCTGCTACGCTGATTCGGTGTGAGCTTCCAGGAATTCATGATCGC 600

QY 651 TCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCGCAACCGCGGAGGTCTACCAAGC 710
 DB 601 TCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCTCGCAACCGCGGAGGTCTACCAAGC 660

QY 711 ACTGAAGTCCGTCATCAAGGAAAGGGGCTCTGTCACCGGACTTTCGCGATGAGGCGGCTT 770
 DB 661 ACTGAAGTCCGTCATCAAGGAAAGGGGCTCTGTCACCGGACTTTCGCGATGAGGCGGCTT 720

QY 771 CGCTCTCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
 DB 721 CGCTCTCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780

QY 831 GGCTGGCTTCACCCAGGCAAGACATCGCTCTTGCTCTGCGAGTGTGCTTCTCTGAGTT 890
 DB 781 GGCTGGCTTCACCCAGGCAAGACATCGCTCTTGCTCTGCGAGTGTGCTTCTCTCTGAGTT 840

QY 891 CTTCAAGGACGGCACTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAA 950
 DB 841 CTTCAAGGACGGCACTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAA 900

QY 951 CGTTTACGCTGAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCCACTTCGAGGA 1010
 DB 901 CGTTTACGCTGAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCCACTTCGAGGA 960

QY 1011 AGATGACTGGAGGGTTTACACCACTTCACCGCAACCATCGGCAACAGGTTCAGATCGT 1070
 DB 961 AGATGACTGGAGGGTTTACACCACTTCACCGCAACCATCGGCAACAGGTTCAGATCGT 1020

QY 1071 TGGCGACGACTTCTTCTGCTACCAACCTGAGCGCTGGAAGGAGGCACTCCCTAAGAAGGC 1130
 DB 1021 TGGCGACGACTTCTTCTGCTACCAACCTGAGCGCTGGAAGGAGGCACTCCCTAAGAAGGC 1080

QY 1131 TGCCTAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCTTCAACCGAGACCTTTCGAGCG 1190
 DB 1081 TGCCTAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCTTCAACCGAGACCTTTCGAGCG 1140

QY 1191 TGTGCACTGGCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAAGC 1250
 DB 1141 TGTGCACTGGCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAAGC 1200

QY 1251 CGAGGACACCACTTGTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
 DB 1201 CGAGGACACCACTTGTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260

QY 1311 TGCTCAGCAGCTTCCGACCGGTGTCGCAAGTAAAGTAAACAGCTTCTCCGCAATCGAGCAGCT 1370
 DB 1261 TGCTCAGCAGCTTCCGACCGGTGTCGCAAGTAAACAGCTTCTCCGCAATCGAGCAGCT 1320

QY 1371 GCTTGGCGACGGCGGTCTACGAGGTGCGAGGCAATTCACGCTTTCAGGGCTTAAT 1430
 DB 1321 GCTTGGCGACGGCGGTCTACGAGGTGCGAGGCAATTCACGCTTTCAGGGCTTAAT 1380

QY 1431 AAAAGCGCTTTTCGAGCG 1448
 DB 1381 AAAAGCGCTTTTCGAGCG 1398

RESULT 10
 BD163286
 LOCUS BD163286 1275 bp DNA linear PAT 17-JAN-2003
 DEFINITION Novel polynucleotide.
 ACCESSION BD163286
 VERSION BD163286.1 GI:27869050
 KEYWORDS JP 2002191370-A/1085.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 1275)
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 1085 09-JUL-2002;
 COMMENT KYOWA HAKKO KOGYO CO LTD
 OS Corynebacterium glutamicum
 PN JP 2002191370-A/1085
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO

```

PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
PC 04, C12P13/08,
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1:15),
PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
PC C12N15/00,
PC C12N5/00, C12N15/00
CC Novel polynucleotide
FH Key location/Qualifiers
FT source 1..1275
FT /organism='Corynebacterium glutamicum'
FEATURES
    source
        Location/Qualifiers
            1..1275
                /organism='unidentified'
                /mol_type='genomic DNA'
                /db_xref='taxon:32644'
ORIGIN
    Query Match 80.7%; Score 1273.4; DB 6; Length 1275;
    Best Local Similarity 99.9%; Pred. No. 2.1e-231;
    Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 151 GTGGCTGAATCATGACGCTATTTCGCTCGCGAAATTCCTCGACTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAATCATGACGCTATTTCGCTCGCGAAATTCCTCGACTCCCGCGGTAAACCAACC 60
QY 211 GTGAGGCGAGAGGTTTCTGTGATGACGTTTCCACGGTGTCCAGGTTGCCAGGTTCCATCCGGC 270
DB 61 GTGAGGCGAGAGGTTTCTGTGATGACGTTTCCACGGTGTCCAGGTTGCCAGGTTCCATCCGGC 120
QY 271 GCATCCACCGGCTCCACAGGCTCATGAGCTCGCGTGAAGGTTGGCGATCGCTACCTGGGC 330
DB 121 GCATCCACCGGCTCCACAGGCTCATGAGCTCGCGTGAAGGTTGGCGATCGCTACCTGGGC 180
QY 331 AAGGCGTTTTGAAGGCGAGTTGAAACGTCACAGGAATTCGGCGACGAGCTCGCTGGC 390
DB 181 AAGGCGTTTTGAAGGCGAGTTGAAACGTCACAGGAATTCGGCGACGAGCTCGCTGGC 240
QY 391 CTAGAGGCTGACGATCAGCGCTCATCGACAGCAATGATCAAGTTGATGGCACCGCC 450
DB 241 CTAGAGGCTGACGATCAGCGCTCATCGACAGCAATGATCAAGTTGATGGCACCGCC 300
QY 451 AACAACTCCGCTGGTGTCAAACGCAATCCTTGTGTTTCCATGGCTTTGCCAAGGCT 510
DB 301 AACAACTCCGCTGGTGTCAAACGCAATCCTTGTGTTTCCATGGCTTTGCCAAGGCT 360
QY 511 GCTGCTGATTCGCGAGGCTCCACCTGTTCCGTCATCGTGGGACCAACGACACGTT 570
DB 361 GCTGCTGATTCGCGAGGCTCCACCTGTTCCGTCATCGTGGGACCAACGACACGTT 420
QY 571 CTTCCAGTTTCCATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTTCAC 630
DB 421 CTTCCAGTTTCCATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTTCAC 480
QY 631 GTTCAAGATTCATGATCGCTCAATCGGTGAGAGACCTTCTTGAAGCTCTCCGCAAC 690
DB 481 GTTCAAGATTCATGATCGCTCAATCGGTGAGAGACCTTCTTGAAGCTCTCCGCAAC 540
QY 691 GCGCGGAGGCTTACACGACGCTGAAGTCCGTCATCAGGAAGGGCTGTCCACCGGA 750
DB 541 GCGCGGAGGCTTACACGACGCTGAAGTCCGTCATCAGGAAGGGCTGTCCACCGGA 600
QY 751 CTTGCGATGAGGGGGCTTCGCTCTTCCGTCGGCTCCACCGTGAGGCTCTTGACCTT 810
DB 601 CTTGCGATGAGGGGGCTTCGCTCTTCCGTCGGCTCCACCGTGAGGCTCTTGACCTT 660
QY 811 ATCGTTAAGGCAATCGAAGGCTGGCTTACCCAGGCAAGGACATCGCTTGTGCTGTG 870
DB 661 ATCGTTAAGGCAATCGAAGGCTGGCTTACCCAGGCAAGGACATCGCTTGTGCTGTG 720

```

```

QY 871 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCACTACCACTTGAAGGTGGCCAGCAC 930
DB 721 GACGTTGCTTCTCTGAGTTCTTCAAGGACGCACTACCACTTGAAGGTGGCCAGCAC 780
QY 931 TCCGAGCTGAGATGCGCAAAAGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 990
DB 781 TCCGAGCTGAGATGCGCAAAAGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 840
QY 991 ATCGAGGACCACTCGCAGGAGATGACTGGGAGGTTTACCAAACTTACCCGCAACCATC 1050
DB 841 ATCGAGGACCACTCGCAGGAGATGACTGGGAGGTTTACCAAACTTACCCGCAACCATC 900
QY 1051 GCGCAAAAGTTTCAGATGTTGGGCGACGACTTCTTCTGTCACCAACCTCGAGCGCTGAAG 1110
DB 901 GCGCAAAAGTTTCAGATGTTGGGCGACGACTTCTTCTGTCACCAACCTCGAGCGCTGAAG 960
QY 1111 GAGGCGATCGTTAAGAGGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCGGTACC 1170
DB 961 GAGGCGATCGTTAAGAGGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCGGTACC 1020
QY 1171 CTCACCGAGACTTCGACGCTGTGACATGCTACCGGCGAGGCTACACCTCCATGATG 1230
DB 1021 CTCACCGAGACTTCGACGCTGTGACATGCTACCGGCGAGGCTACACCTCCATGATG 1080
QY 1231 TCCCAACCGTTCCGTTGAGACCGAGGACACCACTTGTGACTCGCAGTTGCACCTCAAC 1290
DB 1081 TCCCAACCGTTCCGTTGAGACCGAGGACACCACTTGTGACTCGCAGTTGCACCTCAAC 1140
QY 1291 TGTGGCCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGTGTCCGAAAGTACAAACGAG 1350
DB 1141 TGTGGCCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGTGTCCGAAAGTACAAACGAG 1200
QY 1351 CTTCTCCGATCGAGCAGCTGTTGGCGACGCGCGCTTACGAGGTCGCGAGCGCATTC 1410
DB 1201 CTTCTCCGATCGAGCAGCTGTTGGCGACGCGCGCTTACGAGGTCGCGAGCGCATTC 1260
QY 1411 CCAGCGTTTCAGGGC 1425
DB 1261 CCAGCGTTTCAGGGC 1275
RESULT 11
AX121169
LOCUS AX121169 1275 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1085 from Patent EP1108790.
ACCESSION AX121169
VERSION AX121169.1 GI:14037884
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE 1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 1085 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
    source
        1..1275
            /organism="Corynebacterium glutamicum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1718"
ORIGIN
    Query Match 80.7%; Score 1273.4; DB 6; Length 1275;
    Best Local Similarity 99.9%; Pred. No. 2.1e-231;
    Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 151 GTGGCTGAATCATGACGCTATTTCGCTCGCGAAATTCCTCGACTCCCGCGGTAAACCAACC 210
DB 1 GTGGCTGAATCATGACGCTATTTCGCTCGCGAAATTCCTCGACTCCCGCGGTAAACCAACC 60

```

211 GTGAGGCGAGGTTTCTCGATGACGGTTCCCA CGGTGTGCGAGGTGTTCCATCGGC 270
Db
61 GTCGAGCGAGAGGTTTCTCGATGACGGTTCCCA CGGTGTGCGAGGTGTTCCATCGGC 120
Qy GCATCCACGGCGTCCACGAGGCTCATGAGTGGTGACGGTGGCGATCGCTACCTGGGC 330
Db GCATCCACGGCGTCCACGAGGCTCATGAGTGGTGACGGTGGCGATCGCTACCTGGGC 180
Qy AAGGGCGTTTGAAGGCGATTTGAAAACGTCACGAGAAATCGCGACGAGCTCGCTGGC 330
Db AAGGGCGTTTGAAGGCGATTTGAAAACGTCACGAGAAATCGCGACGAGCTCGCTGGC 240
Qy CTAGAGGCTACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGCGACCGCC 300
Db CTAGAGGCTACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGCGACCGCC 300
Qy AACAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCGCTGTGCAAGGCT 510
Db AACAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCGCTGTGCAAGGCT 360
Qy GCTGCTATTCGCGAGGCTCCCACTGTTGCGCTACATCGGTGACCAAAACGACACGTT 570
Db GCTGCTATTCGCGAGGCTCCCACTGTTGCGCTACATCGGTGACCAAAACGACACGTT 420
Qy CTTCCAGTTTCCAATGATGAACATCATCAAGTGGCGCTCA CGCTGACTCGGTTGAC 630
Db CTTCCAGTTTCCAATGATGAACATCATCAAGTGGCGCTCA CGCTGACTCGGTTGAC 480
Qy GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTTGAGGCTCTCCGCAAC 540
Db GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTTGAGGCTCTCCGCAAC 540
Qy GCGCGAGGTTCTACGACGCTCAAGTCCGTATCAAGGAAAAGGCGCTGTCCACCGGA 750
Db GCGCGAGGTTCTACGACGCTCAAGTCCGTATCAAGGAAAAGGCGCTGTCCACCGGA 600
Qy CTTGCGGATGAGGCGGCTTCGCTCTTCCGTGCGCTCCACCGGTGAGGCTCTTGACCTT 810
Db CTTGCGGATGAGGCGGCTTCGCTCTTCCGTGCGCTCCACCGGTGAGGCTCTTGACCTT 660
Qy ATCCTTGGGCAATCGAAGGCTGGCTTCAACCGGCAAGGATCGCTCTTGCTCTG 870
Db ATCCTTGGGCAATCGAAGGCTGGCTTCAACCGGCAAGGATCGCTCTTGCTCTG 720
Qy GACGTTCCTCTGAGTCTTCAAGGCGGACCTTACCACTTCGAGGTGGCGACGAC 930
Db GACGTTCCTCTGAGTCTTCAAGGCGGACCTTACCACTTCGAGGTGGCGACGAC 780
Qy TCCGCGAGTGAAGTGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCCAATCGTCTCC 990
Db TCCGCGAGTGAAGTGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCCCAATCGTCTCC 840
Qy ATCGAGGACCCATCGCAGGAAGATGATGGAGGGTTACCAACCTCACCGGACCATC 1050
Db ATCGAGGACCCATCGCAGGAAGATGATGGAGGGTTACCAACCTCACCGGACCATC 900
Qy GCGGACGAGTTTCAGATCGTTGGGACGACCTTCTTCGTCAACACCTTGAGCGCTGAAG 1110
Db GCGGACGAGTTTCAGATCGTTGGGACGACCTTCTTCGTCAACACCTTGAGCGCTGAAG 960
Qy GAGGCGATCGCTAAGAGGCTGCCAATCCATCTCTGGTTAAGGTGAACAGATCGGTACC 1020
Db GAGGCGATCGCTAAGAGGCTGCCAATCCATCTCTGGTTAAGGTGAACAGATCGGTACC 1020
Qy CTACCGGAGACCTTCGAGCGTGTGACATGGCTACCGCGGAGGTACACCTCCATGATG 1230
Db CTACCGGAGACCTTCGAGCGTGTGACATGGCTACCGCGGAGGTACACCTCCATGATG 1080
Qy TCCCAACGCTTCGGTGAGACCGAGGACCAACCATTTGCTGACCTCGCAGTTGCACTCAAC 1290
Db TCCCAACGCTTCGGTGAGACCGAGGACCAACCATTTGCTGACCTCGCAGTTGCACTCAAC 1140

Qy 1291 TGTGGCCAGATCAAGACTGGTGTCTCGACACGTTCCGACCGTGTGCAAGTACACACG 1350
Db 1141 TGTGGCCAGATCAAGACTGGTGTCTCGACACGTTCCGACCGTGTGCAAGTACACACG 1200
Qy 1351 CTTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTTACGAGGTGCGAGCGCATTC 1410
Db 1201 CTTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTTACGAGGTGCGAGCGCATTC 1260
Qy 1411 CCACGCTTTCAGGCG 1425
Db 1261 CCACGCTTTCAGGCG 1275

RESULT 12

AP005217 300750 bp DNA linear BCT 24-JUL-2003
LOCUS Corynebacterium efficiens YS-314 DNA, complete genome, section 4/11.

ACCESSION AP005217 BA000035

VERSION AP005217.1 GI:23492722

KEYWORDS

SOURCE Corynebacterium efficiens YS-314

ORGANISM Corynebacterium efficiens YS-314

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1 Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E.,

Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and

Gojobori, T.

Comparative complete genome sequence analysis of the amino acid

replacements responsible for the thermostability of Corynebacterium

efficiens

Genome Res. 13 (7), 1572-1579 (2003)

JOURNAL MEDLINE

PUBMED 12840036

REFERENCE 2 (bases 1 to 300750)

Director-General of Biotechnology Center.

Submitted (17-MAY-2002) Director-General of Biotechnology Center,

National Institute of Technology and Evaluation, Biotechnology

Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan

(E-mail: bioinfo.go.jp, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)

Kawarabayasi, Y. is officially affiliated with the National

Institute of Advanced Industrial Science and Technology, Tsukuba,

Ibaraki, 305-8566 Japan

Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the

National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan

Ittoh, T. is at the Japan Biological Information Research Center,

Koto-ku, Tokyo, 135-0064 Japan

Yamagishi, A. is at Tokyo University of Pharmacy and Life Science,

Hachioji, Tokyo, 192-0392 Japan

Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co.,

Inc., Kawasaki, Kanagawa, 210-8681 Japan

The other authors are at the National Institute of Technology and

Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.

FEATURES

source

1. 300750

/organism="Corynebacterium efficiens YS-314"

/mol_type="genomic DNA"

/strain="YS-314"

/db_xref="taxon:196164"

252..620

/notes="CB0885"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="BAC17695.1"

/db_xref="GI:23492723"

/translation="MDQAFSGCLVDCVHGVIYFDVTPVMLLPVVGTSNKRKIVLDLFPN

HYFLNKPDRILFTPLMSGCGCLFGWVIGLTPWHTVYVCVFCGEPDPSGRTLAAC

LTHASRTMKPCLQGLVANG"

529..1936

gene

```

/genes="rRNA_16S"
529..1936
/genes="rRNA_16S"
/product="16S ribosomal RNA"
/notes="CrRNA01"
complement(2042..2215)
/notes="CE0886"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17696.1"
/db_xref="GI:23492724"
/translation="MSSQTPNSMPPTHHHLGDFADPWWMIHPITTCWQHPHGSNT
NPPPHSHSPWVY"
2399..5481
/genes="rRNA_23S"
2399..5481
/genes="rRNA_23S"
/product="23S ribosomal RNA"
/notes="CrRNA02"
5606..5725
/genes="rRNA_5S"
5606..5725
/genes="rRNA_5S"
/product="5S ribosomal RNA"
/notes="CrRNA03"
5925..7088
/notes="CE0889, similar to AX063829-1|CAC25156.1| percent
identity: 79 in 387 aa"
/codon_start=1
/transl_table=11
/product="putative aminotransferase"
/protein_id="BAC17699.1"
/db_xref="GI:23492727"
/translation="MRNKVVERLRPPFGETIIFATMTQRANEAGAINLGGPFDEGPA
PMEIASQILGNQVSGRGLDGLRTAVASDRHRKFLDYDPTEVLITVGATEAI
SASVGLPEGPEVLEPYDAYAAIALAGTAVAVPLKEEDNTWLTDTDAVAAV
TKTSMIIVNSPNTGTSFVSKRSLEALAGIARAYDLVLSDEVVHLTFDKVSHAV
ASLPGMDRTITVSSAAKFTVGTGKTGALAPAPLEAVIRAKOFMSVGVATPQPA
VAHAHAKPWPVKMEKGLONKRDILTDALTQAGLKVDHSHCTYFVADIGERDAEP
CFDILDRGVAAIPVQAFQPEQSSKVRFAFCQEDTLREARQLRAAGSL"
complement(7180..7821)
/notes="CE0890"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17700.1"
/db_xref="GI:23492728"
/translation="MTHFSFDVDEAYARKNNEFLRDTKRLQISAFFFGVLIAIGGVL
YLDAGVGMVNLVIMGMAALSFMIVPLPROVGSASLNTYPLVPVAVINPR
DVLLALVNTNVDLTPPEWGLAVRTINRLGVHDKLGERVPSVAVSGRRTRDQHW
DEISMPITGTTDADVRAEKTIPHLWNKLEKNRNLDEVKTRANLLVL"
complement(7822..9495)
/notes="CE0891, similar to AL583924-93|CAC31112.1| percent
identity: 75 in 547 aa"
/codon_start=1
/transl_table=11
/product="putative DNA helicase"
/protein_id="BAC17701.1"
/db_xref="GI:23492729"
/translation="MAFGDGLIVQSDKTLVLTGHDLAGEARAALAPFAELERAPHE
VHTYRTPLANARAAGHDAQVVDVLEYSRFPVPOPLVDVAETWSYGRVPMKH
HPAHGLLESSEPAILLVELSRHKIKPMLGQQLDDETTAVHPSERGLKQHLKIGWP
ASDLAGYDGEAHPALSTLSEVDSRLDQQAAADSFWEGSGVVVLPCGAGKTWGA
ASMAQAQTLLLVNTVAGROWKDELLRTLTDEIGEYGERKEIRPVYATFYOV
VTRTKGYLLELFDSDRMDGLIYDEVHLLPAPYFRMTSLQSRRLGLTATLVRED
GREGDVFLGPKRDAPKWLDAQOGLATACDCEVRTTMDAERMYATATADRYR
LAATAHTTRVVRVLEHQOQPTLIIGAYLDQELGAEFADPIVQCKTSNKKREL
FDRFSGELNVLVSKVANFSIDLPEAAVAVIOVSGTFGSRQBEAQRGLRLAPKADGG
EAHFSVSVSRDLTDEYAAHRORFLAQGYATRIIDADDLLFPITTEKET"
complement(9548..11851)
/notes="CE0892, similar to AE006976-11|AAK45126.1| percent
identity: 36 in 741 aa"

```

```

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17702.1"
/db_xref="GI:23492730"
/translation="MPCPPFSQIROTKVTSTAPKLGDIIMKSDSPVPTLTGWLQLEDD
QLSGTLNRPTALPLPGLSLAARLQIRASIQRAVHKLSALELAVLEAAAGAEAL
NPITAPFVVERLHRAHRSAAVPTAEQICEALNRLQALVFGGERFIILPTMATLP
TNOLLPEHGRSLTFGEASOIEALPARHKILVLTNSGLGTRDAALGADPADPAP
IPOLIKAGLSRDSQTVKLPGVRRVLEGOEDTAVRVVUPSPSAPGAGDAGIAGAL
EVARLQMLIDALSTAPASTLKEGALGVMTVRLTRELQVDEQQVARVSVLSGSGLL
RNGVDPADDDGDIYAPTVADEWLOLAELQALHMRGWTOTTAAMLVGEADE
RKPTHLSPASRRDALPETRAMILOSLTRVGEEDLEADLFPHPLAASRIPRETIRH
LLEEARWTCVAGGVTSPARVLTQAPAGVPEIETADIAVAVRKPVDYFIQVADYTV
MVPGLQPEMOKIIGQLADLSPLASVYRSEASLRHMDLGLTATEIDFLTOHSA
TGLPOSVALLQDIARRHGTLRGGPALCYLRSDDPALLHSAVEAGALGLRQIATPVA
VSNASLIEVITALKRAGQHPVADLQASLDLAPRPARVPAEQPSRAGPLDESRTIRA
AVAAIRBEVARTGTVSDOPTLAVLQAAVRGORTVTLGFDVKQGVAVHRTVKPLTVNA
GQVDALDESTGAVHRFMLHRITEIVD"
11836..12036
/notes="CE0893, similar to AL583924-91|CAC31110.1| percent
identity: 50 in 62 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17703.1"
/db_xref="GI:23492731"
/translation="MGNWVVEKGYVDPANPDHDPSEGHVVTETLIAPYAGASSPWGD
DMEPFVPAETIGYVHPYTRINR"
complement(12229..12822)
/notes="CE0894, similar to AX133781-1|CAC38957.1| percent
identity: 71 in 195 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC17704.1"
/db_xref="GI:23492732"
/translation="MLFMGRHSAKTATPFVKMAATTVAFGAATAVAFAPAAAPSDW
DLRANCSGNWAINTGNGHGLQFSPSTWAVGGDEFAPPAQATREQQIAVAERT
LAGQGWNPACSKSLGNSAPTRNLNPPAPAPAPAPAPATVDTNTNPPVVGSS
DLNTVQGIYDAVSGTLAQYGVQVPAEQAHYNAFGR"
13342..13725
/notes="CE0895, similar to AE006977-4|AAK45135.1| percent
identity: 52 in 127 aa"
/codon_start=1
/transl_table=11
/product="putative cold shock protein"
/protein_id="BAC17705.1"
/db_xref="GI:23492733"
/translation="MPVGTVMKVDPERGFGVSNPGGDCFCVGKQVLPGKVTTELHOGQ
RIEFDPAGRKGPQALRVKVLTPRRRPQHTYKPEELNGLISDMVTMLEGTVPALLK
GHYDRKVGQVAKILRAVAKLEA"
complement(13881..14408)
/notes="CE0896"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17706.1"
/db_xref="GI:23492734"
/translation="MASRKTGRKFIQIATLLVAVVVVVVSVFQWNNRPRPEPE
TVSTASSPAGEIEVFYPSICEPGVCAENVEPTLEVGADEBLRTIPEIHDHWWL
LTIYDDPAANDEFHTSYETETVTVGSDVPTEGGERPLRVVVVEVTSVMIGTDNGE
ETPITVWSLAVRQ"
complement(14424..15254)
/notes="CE0897, similar to AE001874-1|AAF09704.1| percent
identity: 40 in 267 aa"
/codon_start=1
/transl_table=11
/product="putative glutamine cyclotransferase"

```

Query Match 65.6%; Score 1034.8; DB 1; Length 300750;
 Best Local Similarity 81.1%; Pred. No. 2.7e-186;
 Matches 1234; Conservative 0; Mismatches 247; Indels 41; Gaps 1;

QY 96 CATTGATACAGTGAATTAATCTAGTTAGTCTCCCAAGTTGGCATAGAGGCCACAGTGGC 155
 Db 170262 CCTTTTTCGGGTGAACATATCTGTTAGTTTCCCAAGTTGGCATAGAGGCCACAGTGGC 170321
 QY 156 TGAATCATGACAGTATTCGCTCGGAAATTTCTCGACTCCCGCGGTAAACCAACCGTCCA 215
 Db 170322 TGAATCATGACAGTATTCGCTCGGAAATTTCTCGACTCCCGCGGTAAACCAACCGTCCA 170381
 QY 216 GGCAGAGGTTTCTCGATGACAGGTTTCCACGCTGTGCGAGTGTCTTCATCCGCGGCATC 275
 Db 170382 GGCAGAGGTTTCTCGATGACAGGTTTCCACGCTGTGCGAGTGTCTTCATCCGCGGCATC 170441
 QY 276 CACCGCGTCCACGAGGCTCATAGAGTCTGCTGAGCGGTGGCGATCGCTACTTGGGCAAGG 335
 Db 170442 CACCGGTGTCACGAGGCGCACGAGCTCGTGTGAGCGGCGAGATCGTTACCTTGGGCAAGG 170501
 QY 336 CGTTTGAAGGAGTGAAGAGTCAAGAGAAATCGGCGACGAGCTCGCTGGCCTAGA 395
 Db 170502 TGTTCCTCAAGGCTGTTGAGAACGTTCAAGAGGAGATCGGCGACGAGCTGGCGCGCCTGGA 170561
 QY 396 GGCCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGAATGCGACCGCGCAACAA 455
 Db 170562 GGCCTGACGATCAGCGCTCATCGATGCGGCGATGATCAAGCTTGAATGCGACCGCGCAACAA 170621
 QY 456 GTCCCGCTGGTGCACAAACGCAATCCTTGGTGTTCATGAGTCTGTGCAAGGCTGTGC 515
 Db 170622 GTCCCGCTGGTGCACAAACGCAATCCTTGGTGTTCATGAGTCTGTGCAAGGCTGTGC 170681
 QY 516 TGATTCGCGAGGCTCCGACTGTTCCGCTATCATCGTGGACCAACGACGAGTCTTCTCC 575
 Db 170682 GATTCGCGAGGCTCCGACTGTTCCGCTATCATCGTGGTGGCCGAAACGCGGCTGCTGCC 170741
 QY 576 AGTTTCCATGATGAACATCATCAAGTGGGCTCAGCTGACTCCGCTCGGTGTGACGTTCA 635
 Db 170742 CGTCCGATGATGAACATCATCAAGTGGGCTCAGCTGACTCCGCTCGGTGTGATGTTCA 170801
 QY 636 GGAATTCATGATCGTCCAAATCGGTGTCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGC 695
 Db 170802 GGAGTTCAATGATCGTCCGCTCGGTGTCAGAGACCTTCTCTGAGGCTCTCCGCTCGG 170861
 QY 696 GGAGGTTACACGCACTGAAGTGGTTCATCAAGGAAGGCTGTGTCACCGGACTTGG 755
 Db 170862 GGAGGTTACACGCACTGAAGAGGTTATCAAGGAAGGCTGTGTCACCGGCTCGG 170921
 QY 756 CGATGAGGCGGCTTGGCTCTTCCGCTCGGCTCCACCGCTGAGGCTCTTGACCTTATCGT 815
 Db 170922 CGATGAGGCGGCTTGGCTCTTCCGCTCGGCTCCACCGCTGAGGCTCTTGACCTTATCGT 170981
 QY 816 TAAGGCAATCAGAGAGGCTGCTTCAACCCAGGCAAGGACATCGCTCTTGGAGCT 875
 Db 170982 CGAGGCCATCAAGAGGCGGCTTCCAGGCGGCAAGGACATCGCTCTTGGAGCT 171041
 QY 876 TGCTTCTCTGAGTCTTCAAGGAGGCGACCTTACCACTTCAAGGTCGAGGCTCGGCTCGC 935
 Db 171042 TGCTTCTCTGAGTCTTCAAGGAGGCGACCTTACCACTTCAAGGTCGAGGCTCGGCTCGC 171101
 QY 936 AGCTGAGATGCAAGGTTTACGCTGAGTCTGTTGAGCGGTACCAATCGTCTCCATCGA 995
 Db 171102 CGAGAGATGCGGAACGTTACGCGAGCTGTGTGAGCAGTACCGATCGTCTCCATCGA 171161
 QY 996 GGACCCACTCAGGAGAGTGTGGAGGGTTACCAACCTTCAACCGCAACCATCGGCGCA 1055
 Db 171162 GGATCGGCTGAGGAGGAGCTGGATGGTGTGTGTGCTGACGCGGCGGCTCGGCTG 171221
 QY 1056 CAAGGTTTCAATCGTTGGCGAGGACTTCTTGTGACCAACCTTGAAGGAGG 1115
 Db 171222 CAAGGTTTCAATCGTTGGCGAGGACTTCTTGTGACCAACCTTGAAGGAGG 171281
 QY 1116 CATCGTAAAGAGGCTGCGCACTCCATCCTGTTAAGGTGAACAGATCGGTACCTCCAC 1175
 Db 171282 CATCGTAAAGAGGCTGCGCACTCCATCCTGTTAAGGTGAACAGATCGGTACCTCCAC 171341

QY 1176 CGAGACCTTTCGAGCGCTGTCGATCGCTCACCGCGAGGCTACACCTCCATGATGCCCA 1235
 Db 171342 CGAGACCTTTCGAGCGCTGTCGATCGCTCACCGCGAGGCTACACCTCCATGATGCCCA 171401
 QY 1236 CGCTTCCGCTGAGACCGCGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGG 1295
 Db 171402 CGCTTCCGCTGAGACCGCGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGG 171461
 QY 1296 CCAGATCAAGACTGTGTGCTCCAGACGTTTCCGACCGTTCGACCGTTCGCAAAAGTCAACAGCTTCT 1355
 Db 171462 CCAGATCAAGACTGTGTGCTCCAGACGTTTCCGACCGTTCGACCGTTCGCAAAAGTCAACAGCTTCT 171521
 QY 1356 CGCATCGAGGAGCTGTGTGCGAGCGCGGCTTACGAGGTCGAGCGCATTCGCCACG 1415
 Db 171522 CGCATCGAGGAGCTGTGTGCGAGCGCGGCTTACGAGGTCGAGCGCATTCGCCACG 171581
 QY 1416 CTTTCAGGCGCTAAA-----TAAAA 1434
 Db 171582 TTTTCAGGCGCTAGACGCTTTTCGAGAAACCTTGACCCCGCATCCCGTCACTCTCCAAA 171641
 QY 1435 GCGCTTTTTCGACCGCGGTAACCTCAAGGTTGCGGCGGCTGTTGCTTACTACTGTTC 1494
 Db 171642 AGGTGCGGAGTGTGCGGTATCGGGCATGTTGCGGCGTTCGCAAGAGCAGCTGCGGTGAC 171701
 QY 1495 TGGTGTGACTATGATCGAGGTTATGCAAGCAGAGGAACTCATAAAGGCTTGTTC 1554
 Db 171702 GGGTGGGACTATGATCGGTTATGCGGAAGCAGAGAAACCCACAGGGGCTTGTTC 171761
 QY 1555 CTGCTCAAGCAGGGAACGTGC 1576
 Db 171762 CGGTCTCCAGCAGATCGTGC 171783

RESULT 13

LOCUS BX248356 347625 bp DNA linear BCT 06-NOV-2003
 DEFINITION Corynebacterium diphtheriae gravis NCTC13129, complete genome;
 segment 3/8.
 ACCESSION BX248356 BX248353
 VERSION BX248356.1 GI:38199583
 KEYWORDS complete genome.
 SOURCE Corynebacterium diphtheriae
 ORGANISM Corynebacterium diphtheriae
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 (bases 1 to 347625)
 AUTHORS Cordero-Tarraga, A.M., Efstratiou, A., Dover, L.G., Holden, M.T.G.,
 Pallen, M., Bentley, S.D., Beara, G.S., Churcher, C., James, K.D., De
 Zoysa, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T.,
 Hamlin, N., Holtroyd, S., Jørgensen, K., Moule, S., Quail, M.A.,
 Rabinowitz, E., Rutherford, K., Thomson, N.R., Unwin, L.,
 Whitehead, S. and Barrell, B.G. Parkhill, J.
 The complete genome sequence and analysis of Corynebacterium
 diphtheriae NCTC13129
 Nucleic Acids Res. 31 (22), 6516-6523 (2003)
 JOURNAL PUBMED 14602910
 REFERENCE 2 (bases 1 to 347625)
 AUTHORS Cordero-Tarraga, A.M.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2003) Cordero-Tarraga A.M., submitted on behalf
 of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust
 Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:
 anct@sanger.ac.uk
 FEATURES
 source
 1. 347625
 Location/Qualifiers
 /organism="Corynebacterium diphtheriae"
 /mol_type="genomic DNA"
 /strain="NCTC13129"
 /db_xref="taxon:1717"
 /note="biotype gravis"
 complement(81..206)
 /locus_tag="DIP0718"
 complement(81..206)
 gene
 CDS

```

/misc_feature
/locus_tag="DIP0718"
/notes="Doubtful CDS. No significant database matches"
/codon_start=1
/transl_table=11
/product="Putative membrane protein"
/db_xref="GI:38199584"
/translation="MKYRSHNIIVCSLIPPQPLKARHIPAKIPSSPPVRIVELP"
complement(141..206)
/locus_tag="DIP0718"
/notes="Signal peptide predicted for DIP0718 by SignalP 2.0
HMM (signal peptide probability 0.699) with cleavage site
probability 0.667 between residues 22 and 23;
signal-peptide site"
257..484
/locus_tag="DIP0719"
257..484
/locus_tag="DIP0719"
/locus_tag="DIP0719"
/notes="Similar to Mycobacterium tuberculosis CDC1551
hypothetical 10.1 kDa protein MT3304 TR:AAK47646
(EMBL:AE007142) (97 aa) fasta scores: E(): 4.5e-11, 51.38%
id in 72 aa"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49241.1"
/db_xref="GI:38199585"
/translation="WDIKIGFSDSPRELVSRSREHNEIVRIHNRALRDAEGVLDDVD
DGNRYLVNRIAYVEVGSTAPRTVGFAGA"
487..1365
/locus_tag="DIP0720"
487..1365
/locus_tag="DIP0720"
/notes="Similar to Mycobacterium tuberculosis hypothetical
31.0 kDa protein RV3207c or MTCY07D11.19 TR:005859
(EMBL:Z95120) (285 aa) fasta scores: E(): 1.7e-44, 46.6%
id in 287 aa"
/codon_start=1
/transl_table=11
/product="Putative membrane protein"
/protein_id="CAE49242.1"
/db_xref="GI:38199586"
/translation="MTDPAESFFVRFARDYGRAYAIPLVAVITVWLVLDVFRTPAE
TTTTATGCAPTATATACGKGPDPARONRPAIATLPGSPFTQKGRGTVRTVGN
GAHAGKHDPKPTVLEVGINTAYAGGDDAFAMVDATLNPKSWTHDKPGPEHV
DAGAVKDPRLQLSDVDTGHLGCGNIAMETSCFIGNRVVINESRWVRGAKFPQG
DIGAVRQYLINHEVHGIGIFANHEPCGKNGELAPIMMOQTLSLSNELFAIDANETYN
DGCVCASNPWPYPFA"
535..603
/locus_tag="DIP0720"
/notes="1 probable transmembrane helix predicted for
DIP0720 by TMHMM2.0"
1165..1194
/locus_tag="DIP0720"
/notes="ScanRegExp hit to PS00142, Neutral zinc
metallopeptidases, zinc-binding region signature."
1389..2246
/locus_tag="DIP0721"
1389..2246
/locus_tag="DIP0721"
/notes="Similar to Mycobacterium leprae hypothetical
protein ML0818 TR:Q9CGG7 (EMBL:AL583919) (297 aa) fasta
scores: E(): 1.2e-34, 39.47% id in 266 aa"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49243.1"
/db_xref="GI:38199587"
/translation="MSLQQLQVHRDAFHPNTPPQLGVANDYGRVGNVAFPSQVIH
PRMSASARVRLKQPGQVVRVPLKSTDGRYINAGWRANSFEEGLARRVDETVAAL
LRDLTALADIVDFSPEDSHLPSLADHSAWAPMVVALGVTSDAARVETPQAL
AQRLLPKVILLRDLIDAPIQVTHADMFAFTIYSGNOMPVVTLGVARYPYGTAACVI
VDALLAGAVDEGIIDRFSHIQRDQLLALVYRIAVHALHPESISNTGNLEWISQT

```

```

IMSRASVTL"
2250..5432
/locus_tag="DIP0722"
2250..5432
/locus_tag="DIP0722"
/notes="Similar to Streptomyces coelicolor putative
ATP-dependent DNA helicase 2SC3B6.07 TR:Q9FCK5
(EMBL:AL390968) (1159 aa) fasta scores: E(): 6.2e-29,
32.21% id in 1161 aa"
/codon_start=1
/transl_table=11
/product="Putative ATP-dependent DNA helicase"
/protein_id="CAE49244.1"
/db_xref="GI:38199588"
/translation="MDAFSTPDQSPPTSAQVIGDLTAGIPLPQOSFVYQAPRVEIV
YNDHSEPIREWDNQLPQORGTWRVTGAGAGVSTLLMDTVAERTRGVSPSIIIVA
ASRESAARLRAGIAHRIADGTSYPASLVRSVHSLAFALRSLSDQURLIITGAEOGA
VIRELLQGHVDNPSLAANPAEQALGLIGAREVDRDFLLSRGRLGEPDELELGA
RYGRPMWASAGKFMREYQQVMNLGTHSLNASLSEKLCFLDIPMGWHTIIDDAQH
LDPOSAQLQALMRYTDFVTAGDPQQSFHFRGASPEFLHAPDHRTLYSSPROP
TVEAKVIGSTGDQSEFVADLLRRSHLLRGIRWDMNAVIRVADIPSLRRALLSAGVP
VOEDPSDILSEQRIVSALILAVRAIQELSNAQLEELALGPICGADTTLRLFRGL
RKAENHAGNRRAIEMRELIDPKESESQTLREQVEAVLTDRELAVDVKIRAVLRG
AOPGSEVIELWEIWDASLSLQTVSLRGVGRGAORDDLDAVMALFDAGDGVERR
PTASITSEVRHIAEQELPTGVRRLATADAVRIVTAHGSQQQWHTVAVGQGTW
PSLOETGLTGOEOLIDLIDSGIEPTVYISRAEKLKEKRLFHVTRGTGRVTVTA
VESPEDTAARSPFLOQIASEKVGQPTTDPETASVQSEIDTETVHLDOYVLLSVPS
IVAEIRRLANPDSFORREQAARQLARLALHQPVGARPEQWVGGSSETSLDISK
VPSLIENALICPLRLARLERLVEEENTPIHMLKGTLAHAFABEAVGRGVDPNAEQLVT
QAFEALLDAPASLPHHMSOWSTMQLRAHMDVSHAQRLVGVGVVNVVAVGVEL
RGRIDRLERNDAEFHIVDFKTGQAVTKDBANENKQLAYQLALHRLGKLMORNGEPA
INTTAEOPLGTVDOAVLYPATDKTIVTTREQAPKDAEELKFSATLPALLSLSGPQ
LVARINFCDOCKIKTMCPAQPEGKWPEC"
2250..2558
/locus_tag="DIP0722"
/notes="ScanRegExp hit to PS00430, TonB-dependent receptor
proteins signature 1."
3045..3200
/locus_tag="DIP0722"
/notes="HMMpfam hit to PF00580, UvrD/REP helicase"
3285..3431
/locus_tag="DIP0722"
/notes="HMMpfam hit to PF00580, UvrD/REP helicase"
4167..4199
/locus_tag="DIP0722"
/notes="ScanRegExp hit to PS00136, Serine proteases,
subtilase family, aspartic acid active site."
5426..8656
/locus_tag="DIP0723"
5426..8656
/locus_tag="DIP0723"
/notes="Similar to Mycobacterium tuberculosis CDC1551
helicase, UvrD/Rep family MT3295 TR:AAK47638
(EMBL:AE007142) (1101 aa) fasta scores: E(): 2.9e-25,
43.03% id in 1120 aa"
/codon_start=1
/transl_table=11
/product="Putative helicase"
/protein_id="CAE49245.1"
/db_xref="GI:38199589"
/translation="MLSPQELSCALGOKFPPTPOQADVISSPLAPLVVAGAGAKTE
TMAARVNLVAGLVDPDRVLGLTPRKAQQLSKRDLRLEQLAGINLRDLPTGA
LATKLEAIPVTSTYDSAGRLIKEYGLLPFPSSRLISQTLFQIAHSIVSAHTGA
LNTSNEATVTSTLISLVSENMHNSPTDIEESAAFLAMIEDVATSKRAPSKEVY
KWRDQTVLRNELLPLVQLKTHLADNHLMTFGEQMSLAARAAENPQVGAQRNRYOI
IMLDEYODTGHQAVLLKSLFAGTAVTAGVPMQSIYGRGATAANLERFLTDFGNSG
SPATKELTVSFRNPPVELDIANFVSRELLGVPRPQVPLEPSPALANGIVRLGF
PPSNDERSYVADHLAAQYEGHDGHPPTAAVIRKRXHSAIALELQORGVPEVIG
LGLGIPVADLVIAITALLVRDPTQAMRILAGSPVLGMADLMALDRAYLSNGR
DRRTATLSRLDRLEKQIITADTTPSQDSIVTGLAEADVADLERDSDGPRYSAGKS
ERLTAAARLYRLNTSLNSLPDLFADIERVFGIRTEVFLQDRSDSGATGTAKHDLR
FAEVQDFSRIPGANLSLLDYLALAESEENGLPEGEVQVTADEVQIILTVHAKGLEW
QHVAVLHADANTYVAKASTWLTNANASVPSALRGADKGDDELVGAPVFEDTPTAAEL

```


AKAGAHIAFDKQVAEENARLFVVAITRAEQVVLVTASADPSKRPVLPYELTMLR	
NDPDSVBERHEGEADYVPPAQEAVFPFNIVVGAEDVFAAMQKQDPLISDDDLF	
ERMEKVSALITEEHOQSAPVAVNIGVELTATDINLAKNPENFAQRERPPVFFKPN	
SYAKGTAFHEMLNRFCAEALDELTELPGLIBELDDLDRLKEAFPLDSEWADRTPE	
HVEHPFVSLGRHIVGRMDAVFNKSDGRWIVVDWTKGPPTDNHKEKESVAMQLAVR	
LAWARLAGIDVDRVDAVPHYVGRNVTYRPPQLPDGBELALKLDPVS"	
5474. .6931	
/locus tag="DIP0723"	
/note="HMPfam hit to PF00580, UvrD/REP helicase"	
5531. .5554	
/locus tag="DIP0723"	
/note="ScanRegExp hit to PS00017, ATP/GTP-binding site	
motif A (P-loop)." .	
8701. .9789	
/locus tag="DIP0724"	
8701. .9789	
Query Match	58.0%; Score 915.8; DB 1; Length 347625;
Best Local Similarity	81.4%; Pred. No. 1e-163;
Matches 1061; Conservative	0; Mismatches 242; Indels 0; Gaps 0;
QY 147 CACAGTGCTGAATCATGACGATTCGCTCGCGAAATTCCTGACTCCCGCGTAACCC	206
DB 201513 CACAGTGCTGACATTTAGCAGTATTTGCTCGTGAATCTTGACTCCCGCGTAACCC	201572
QY 207 AACCGTCAGGCGAGAGGTTTTCTGGATGACGCTTCCACGCGTTCGACAGTGTTCATC	266
DB 201573 AACCGTTGAAGCGAAGTTTTCTTGGATGACGCTTCCACGCGTTCGCGCGTTCCTTC	201632
QY 267 CGGCGCATCCACCGCGTCCACGAGGCTCATGAGCTGCGTGACGCGTGCGATCGCTACCT	326
DB 201633 CGGTCGCTCCACCGGTGTTTCAAGAGCTTCAAGAGCTTCCGACGCGTGCGAGCGCTACCT	201692
QY 327 GGGCAAGGCGTTTTGAAGGCGATTTGAAAAGCTCAAGCAAGAAATCGCGCAGGCTCGC	386
DB 201693 CGGCAAGGCGGTTCTCAATGAGTTAAACAGTGAAACGAGGAATCGCTGACGCAATCGC	201752
QY 387 TGGCCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGGCAC	446
DB 201753 TGGTCAGAGCGGACGATCAGCGCTGATCGATCAGCGGATGATGCTCTCGACGGCAC	201812
QY 447 CGCCAAAGTCCCGCGCTGGGTGCAACGCAATCTTGGTGTTCCTTCCATGGCTGTGCAAA	506
DB 201813 TGAGAAACAGTCTCGTCTCGCGCAACGCTATCTTGGGTGTATCTCATCGCGCTAGCTAA	201872
QY 507 GGCTGCTGATTCGCGACGCTCCACTGTTCCGCTACATCGGTGACCAAAAGCGACA	566
DB 201873 GGCTCGCGAGATCTGCTGGCCTTACCTTTGTACCGCTACATCGCGCGCCTTAAGCTCA	201932
QY 567 GGTCTTCCAGTTCCTCAATGATGAACATCATCAACGCTGGCGCTCACGCTGACTCGGTGT	626
DB 201933 GGTCTTCCAGTTCCTATGATGAACATTTGTTAAACGCTGGCGCACAGCTGACTCGGGGT	201992
QY 627 TGACGTTTCAAGAAATTCATGATCGCTTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCG	686
DB 201993 TGATGTTTCAAGAGTTTCATGATGCTTCTTATCGGTGCGAGTCTTCTCTGAGGCTCTCG	202052
QY 687 CAACGCGCGAGGCTTACCGACGCTCAAGTCCGCTCATCAAGAAAAGGCGCTGTCCAC	746
DB 202053 CATGGGTGACAGGCTTACCACTCTCTTGAAGTCCGCTGATTAAGTCAAGGAGCTTTCAC	202112
QY 747 CGGACTTGGCGATGAGGCGGCTTCGCTCTTCCGCTCCAGCTCCAGCTGAGGCTCTTGA	806
DB 202113 CGGCTTCGGCGAGAGGTGGTTTTCGACCTTCTGTTAGTCCACCAAGGCGAGCTCTGA	202172
QY 807 CTTTATCGTTAAGGCAATCGAGAAGCTGGCTTCAACCCAGCAAGGACATCGCTCTTGC	866
DB 202173 CCTCATCGTTGAGGCAATTTGAGAAGCTGGCTTCAAGCAGGTGCTGACATCGCACTTGC	202232
QY 867 TCTGACGCTTCTCTCTGAGTCTTCAAGGACGCGCACTTCACTTCTGAGGTGGCCA	926
DB 202233 ACTCGAGTGTGCTTCTCTCGAGTTCTTCAAGGATGCGAAGTACCACTTCTGAGGGGGCGA	202292

QY 927 GCACCTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGAGCGCTACCCATCGT	986
DB 202293 GCACACCGCTGAGGAGATGGCAAGGCTTACGAGGAGCTCATCGCTGAGTACCCAATTGT	202352
QY 987 CTCCTATCGAGGAGCCCACTCGACGAAGATGACTGGGAGGGTTACACCAACTCACCGCAAC	1046
DB 202353 TTCCATCGAGGAGCCCACTCGACGAAGAGGACTGGGAGGGCTACACCGCTGACCGCGC	202412
QY 1047 CATCGCGCAAAAGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACCTCGAGCGCT	1106
DB 202413 AATCGGTGACAAGGTTTCAGATCGTTCGCGACGACTTCTTCTGTCACCAACCCAGCAGCT	202472
QY 1107 GAGGAGGCGATCGCTTAAGAGCTGCCAATCCATCCATCTGTTTAAAGTGAACAGATCGG	1166
DB 202473 CAAGGAAGGATCGAGAAGAGGCTGCCAATGCTTCTGCTGGTAAAGGTCAACAGATCGG	202532
QY 1167 TACCCTCACCGAGACCTTCGACGCTGTCGACATGGCTCACCGCGCAGGCTTACACCTCCAT	1226
DB 202533 TACCCTGACCGAGACCTTCGACGCTGTTGATCTCGACACCGCAACGGCTACCGCACCAT	202592
QY 1227 GATGTCCTCCACCGTTCGCGTGAGACCGAGGACACCAATTCGTGACTCGAGTTGACT	1286
DB 202593 GATGTCCTCCACCGCTTCGCGTGAGACTGAAGACACCAATTCGTGACTTCGAGTGCAT	202652
QY 1287 CAACTGTGGCAGATCAAGTCTGCTCCGACGCTTCGACCGGTTCGCGAGGTGCGAAGTACAA	1346
DB 202653 GGGCTGTGGCAGATCAAGACCGGTGCACGCTGCTTCGAGCGGTTCGCAAGTACAA	202712
QY 1347 CAGCTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTCTAGCGAGGTGCGACGCG	1406
DB 202713 TGAGCTTCTCGGATCGAGCAGCAGCTTGAGATGACGAGTCTACGAGGCGGTTTCAGC	202772
QY 1407 ATTCCACGCTTTTCAGGGCTTAAATAAAGCGCTTTTCGACGCG 1449	
DB 202773 ATTTCCAGGTTTTTCAGGGCTTAAATAAAGCTGTAATTAACCC 202815	
RESULT 14	
SCO939115/c	
LOCUS	276800 bp DNA linear BCT 11-FEB-2003
DEFINITION	Streptomyces coelicolor A3(2) complete genome; segment 12/29.
ACCESSION	AL939115 AL035540 AL035554 AL035707 AL049573 AL132674 AL354048
AL355832 AL359989 AL442120 AL596138 AL645882	
AL939115.1 GI:24426505	
VERSION	
KEYWORDS	Streptomyces coelicolor A3(2)
SOURCE	Streptomyces coelicolor A3(2)
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE	
AUTHORS	1 Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornby, T., Howarth, S., Huang, C.H., Kleser, T., Larke, J., Murphy, L., Oliver, K., O'Neill, S., Rabinowitz, B., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrall, B.G., Parkhill, J. and Hopwood, D.A.
TITLE	Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)
JOURNAL	Nature 417 (6885), 141-147 (2002)
MEDLINE	21996410
PUBMED	12000953
REFERENCE	2 (bases 1 to 276800)
AUTHORS	Bentley, S.D.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAY-2002) Submitted on behalf of the Streptomyces sequencing team, Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: sdb@sanger.ac.uk
COMMENT	On or before Oct 30, 2002 this sequence version replaced gi:4500374, gi:4490616, gi:4490978, gi:6138834, gi:7672242, gi:7799503, gi:6894754, gi:10241774.

FEATURES

source	Location/Qualifiers	misc_feature
stem_loop	1. .276800 /organism="Streptomyces coelicolor A3(2)" /mol_type="genomic DNA" /strain="A3(2)" /db_xref="taxon:100226" 48. .98 /note="Score 59: 21/22 (95%) matches, 0 gaps" complement(146. .652) /gene="SCO3054" /note="synonym: SCBAC19G2.09c" complement(146. .652) /gene="SCO3054" /note="SCBAC19G2.09c, conserved hypothetical protein, len: 168aa; similar to many in Streptomyces coelicolor eg. TR:Q9F2X0 (EMBL:AL392148) SCD20.09 hypothetical protein (190 aa) fasta scores; opt: 250, Z-score: 285.6, 34.503% identity (38.816% ungapped) in 171 aa overlap." /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44519.1" /db_xref="GI:14970940" /db_xref="UniProt/TREMBL:Q93J49" /translation="MSRSPAPGGLGVLTVTLTDLTGADSLDTEGRAPPGLTED DYPARELRERATLLAHGHPHRAVTPFLGELLAAAPLVVAVDPADCSAALAPAGE GSLTARVAQAALVAGTWSRLKACAEADCHWAYVDRSPAGRGWCSNQVCCARAKM RYRRAKES" 764. .1123 /gene="SCO3055" /note="synonym: SCBAC19G2.10" 764. .1123 /gene="SCO3055" /note="SCBAC19G2.10, conserved hypothetical protein, len: 119aa; similar to many in Streptomyces coelicolor e.g. TR:O85701 (EMBL:AF072709) 3SCF60.11c conserved hypothetical protein (131 aa) fasta scores; opt: 322, Z-score: 387.9, 42.017% identity in 119 aa overlap and downstream neighbouring CDS SCBAC19G2.11" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44520.1" /db_xref="GI:14970941" /db_xref="UniProt/TREMBL:Q93J48" /translation="MPVIELGVPVLDLCPDPRALARFYADVLGTVGEGDWDRLRPG GRALAPQAPGVPVPRKWPAPDDSQFHLDLVDKLDAAEKAVLELGAERPLDABDRGTG FRVYADPAGHPFLCAC" 786. .1114 /gene="SCO3055" /note="directly repeated at 11832. .12181" 1153. .1545 /gene="SCO3056" /note="synonym: SCBAC19G2.11" 1153. .1545 /gene="SCO3056" /note="SCBAC19G2.11, conserved hypothetical protein, len: aa: similar to many in Streptomyces coelicolor eg. TR:Q9F3E9 (EMBL:AL450289) SC2H2.06 hypothetical protein (150 aa) fasta scores; opt: 340, Z-score: 415.4, 46.032% identity in 126 aa overlap and upstream neighbouring CDS SCBAC19G2.10" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44521.1" /db_xref="GI:14970942" /db_xref="UniProt/TREMBL:Q93J47" /translation="WAPARFRSVVDCPDPRALARFYAAGCGTPEADPDWVVLQV PGPRLSFORAPDLTPPEWFRSDRNAQQHFLDFDGGATWAEMDAADRVLALGARPLD LEDREDKQFVADPAGHPFLCLRIET" 1178. .1527 /gene="SCO3056"	gene
gene	/note="Directly repeated at 11440. .11768" complement(1558. .2688) /gene="SCO3057" /note="synonym: SCBAC19G2.12c" complement(1558. .2688) /gene="SCO3057" /note="SCBAC19G2.12c, possible dipeptidase, len: 376 aa; similar to many eukaryotic peptidases e.g. SW:P31430 (MDPI RAT) rat renal dipeptidase (410 aa) fasta scores; opt: 440, Z-score: 474.0, 36.056% identity (38.906% ungapped) in 355 aa overlap. Also weakly similar to TR:Q9KH70 (EMBL:AF268476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; opt: 273, Z-score: 298.4, 28.159% identity (30.116% ungapped) in 277 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptidase, Renal dipeptidase. Also similar to neighbouring upstream CDS SCBAC19G2.13c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44522.1" /db_xref="GI:14970943" /db_xref="GOA:Q93J46" /db_xref="UniProt/TREMBL:Q93J46" /translation="MADIQDLHTGTEAGLDDLFPVLAEEAFPPPEYAPVDPDDPEP LARAHVAHPADGACNGLPWLRHLPMYDLELCEASAVDTDPRLREHGVALLWSL HLPRLDGDRAVGATLEQLDLVKTVPRAHPGLRLAYDAGAI DARNCGRIAVLPGPA GAALGDELGLRSILHGLRLVLTLSGVSWASEAGLTFRGEVVRMRLGVVADLSG ASATVTRTPAVSKAPALCTRSARALRPHANLPDDLVELGAAGGCMVPTTAEQT GPTVRVADHLDHVRTVAGPQSVGLSYDGAHPLELGPSPCYPRLLVAELLRRGWD RADVALLTGNVQVRVLAFAAFTAKAAQLRRREFSTATIADLDQ" complement(1561. .2616) /gene="SCO3057" /note="Pfam match to entry PF01244 Renal dipeptidase, Renal dipeptidase, score 121.70, E-value 1.4e-32" complement(1562. .2643) /note="Degenerately, directly repeated at 13490. .14610" complement(2808. .4010) /gene="SCO3058" /note="synonym: SCBAC19G2.13c" complement(2808. .4010) /gene="SCO3058" /note="SCBAC19G2.13c, possible dipeptidase, len: 400 aa; similar to many eukaryotic peptidases e.g. SW:P31429 (MDPI_RABIT) rabbit renal dipeptidase (410 aa) fasta scores; opt: 788, Z-score: 879.1, 42.432% identity (48.580% ungapped) in 403 aa overlap. Also weakly similar to TR:Q9KH70 (EMBL:AF268476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; opt: 450, Z-score: 506.1, 31.034% identity (36.486% ungapped) in 348 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptidase, Renal dipeptidase. Also similar to neighbouring downstream CDS SCBAC19G2.12c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44523.1" /db_xref="GI:14970944" /db_xref="GOA:Q93J45" /db_xref="UniProt/TREMBL:Q93J45" /translation="MTSLKARLLRFPVVDHNDLPWALREOVRYDLDARDIAADQ SAHLHTLARLSGGVGAQYMSVVRSDLPGAVTATLQIDCVRLDRHPGELRAAL TAADMEARAEGRIASLMGACGGHSINSLATLALYALGVRYWTLTHNDNADSA TDFPGVGLSAPFREVVRMREGMLVDLSHVAATTHRDADTSTAPVIFSHSSRAV CDHPNIPDDVLERLSANGGMVTFVFKVLAADVMTAADNMRAGHFHLDSSP EAMKHAAPFERVPVATVTADHLDHREAVAGVHILGGYDGTPTPDGLGDV SGYNLITAEILLDRGWSQSLAKLTWKNARVLDAAEDVSRGLRAAGSPSNATIEQLDG TAASQPSG" complement(2820. .4010) /gene="SCO3058" /note="Pfam match to entry PF01244 Renal dipeptidase, Renal dipeptidase, score 280.40, E-value 2.3e-80" complement(2836. .3956)	repeat_region
repeat_unit		misc_feature

```

/notes="Degenerately, directly repeated at 12216..13297"
complement(4025..4567)
/genes="SCO3059"
/notes="synonyms: purE, SCBAC19G2.14c"
complement(4025..4567)
/genes="SCO3059"
/BC number="4.1.1.21"
/notes="SCBAC19G2.14c, phosphoribosylaminoimidazole
carboxylase catalytic subunit PurE, len: 180aa: strongly
similar to many eg. SW:Q44679 (PurE, CORAM)
phosphoribosylaminoimidazole carboxylase catalytic subunit
PurE from Corynebacterium ammoniagenes (177 aa) fasta
scores: opt: 666, z-score: 735.8, 65.541% identity
(65.541% ungapped) in 148 aa overlap and TR:080937
(EMBL:AC004684) putative phosphoribosylaminoimidazole
carboxylase from Arabidopsis thaliana (645 aa) fasta
scores: opt: 644, z-score: 703.2, 60.870% identity
(60.870% ungapped) in 161 aa overlap. Contains Pfam match
to entry PF00731 AIRC, AIR carboxylase."
/codon_start=1
Query Match 46.4%; Score 732.4; DB 1; Length 276800;
Best Local Similarity 72.5%; Pred. No. 6.4e-129;
Matches 962; Conservative 0; Mismatches 361; Indels 3; Gaps 1;

QY 106 CGTGAATAATCTAGTTAGTCCCAAGTTCGCATAGGAGCCACAGTGGCTGAAATCATG 165
DB 43666 CGTGAAGCGTCACAGTGAAGCGTCACATTAAGAAGGAGATGCTGTCGCTCATCGAC 43607

QY 166 CACGATATTCGCTCCGGAATTCGCACTCCCGCGTAAACCAACCGTCGAGGAGAGTT 225
DB 43606 GTGCTGTAGCCCGGGAATCTCTGACTCCCGAGGCAACCCACGTCGAGGTCGAGT 43547

QY 226 TTCTCGATAGCGTTTCCACGCTGTCGAGTGTTCATCCATCCGCGCATCCACGCGTCC 285
DB 43546 GGCCTCGACAGCGCAGCACGGTGTGTCGCGCCCTTCCTGTCGCGGCTCCACCGCGCC 43487

QY 286 CACGAGCTATGAGCTGCGTACGCTGGC---GATCGTACCTGGGCAAGGGGCTTTTG 342
DB 43486 TTCGAGGCCATCGAGCTGCGTGAACGCGGACCCGAGCGCTTACCTCGGCAAGGGGCTCGAG 43427

QY 343 AAGCAGTTGAAACGTCACGAAGAATCGGCAACGAGCTCGCTGGCCTAGAGGCTGAC 402
DB 43426 AAGCCGTGTCGGCGTCTATCGACAGATCGGCCCGGAGCTGTCGCTACGAGCGCCACC 43367

QY 403 GATCAGCGCTCATCGACGAAGCAATGATCAAGTTGATGCGACCGCAACAGTCCCGC 462
DB 43366 GAGCAGCGCTGATCGACAGCGCATGTTTCGACCTGAGCGCCACCGCAACAGGGCTCG 43307

QY 463 CTGGTGAACGCAATCTTGGTGTTCATGCTGTTGCAAGGCTGCTGATTC 522
DB 43306 CTCGCGCAACGCGCATCTCGGGTGTCTCTCGCGCTGCGCCACGCGCGCTCCGAGGCC 43247

QY 523 CGAGCGCTCCACATGTTCCGCTACATCGGTGAGCAACAAACGACACGTTTCTTCCAGTTCCA 582
DB 43246 AGCGACCTGCGCTCTTCGCTACTCTGGCGGCGCCGACGCGACCTGCTGCGGTGCGG 43187

QY 583 ATGATGAACATCATCAACGCTGGGCTCACTGCTGAGCTCGGTGTGAGCTTCAGGAATTC 642
DB 43186 ATGATGAACATCTGAACGGGGCTCGCAACGCGGCTCAACAGTGGACATCCAGGAGTTC 43127

QY 643 ATGATGCTCAATCGGTGAGAGACTTCTGAGGCTCTCGCAACGCGCGCGAGGTC 702
DB 43126 ATGATGCGCCCGATCGGCGGAGTCTCTCTCGAGGCGTCTGCTGGGAGCGCGAGTTC 43067

QY 703 TACCACGCACTGAAGTCCGCTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAG 762
DB 43066 TACCACACCTCAAGAGTCTCTGAAGAACAGGCGCTGGCCACCGCGCTCGGCGACGAG 43007

QY 763 GCGCGCTCGCTTCCTTCGCTGCGCTCCACCGGTGAGGCTCTTGACCTTATGTTAAGCA 822
DB 43006 GCGCGCTTCGCGCCGAACTGGGCTCCAAACCGGAGGCGCTCGACCTCATCTCGAGGCG 42947

```

```

QY 823 ATCAGAGAGGTGCTTCACTCCAGGCAAGACATCGCTCTTGTCTTGGAGCTGCTTCC 882
DB 42946 ATCAAGGAGGCGGCTACACCCCGGCGAGCAGATCGCCCTCGCGCTCGACGTGCGCGC 42887

QY 883 TCTGAGTTCTTCAAGGAGCGGACCTACCACTTCGAAAGTGGCCAGCACTCCGCGAGTCCAG 942
DB 42886 TCCGAGTTCTAAGGAGCGGCTCTTACGCTTCGAGGGCAAGAACCGCTCCGCGCGCGAG 42827

QY 943 ATGCAAAAGCTTACGCTGAGCTGTTGACGCGTACCCAAATCGTCTTCCATCGAGACCCA 1002
DB 42826 ATGACCGAGTACTACGCGAGCTGGTCGAGGGTACCCGCTGGTCTCCATCGAGACCCG 42767

QY 1003 CTGCAAGGAGTGAATCTGGAGGGTTACCAAACTCAGCGCAACCAATCGGCGAGAGTT 1062
DB 42766 CTGTTTCGAGGAGCGACTGCGGAGGGTGGAAACCACTCACCGCAAGCTCGGCGCAAGGTG 42707

QY 1063 CAGATCGTTGGGAGCGACTTCTTGTGTCACCAACCTGAGCGCTGGAAGGAGGCGATCGCT 1122
DB 42706 CAGCTGGTCCGCGAGCGACTGTTGCTGTCACCAACCGGAGCGCTGCGCGCGGATCGAG 42647

QY 1123 AAGAAAGGCTGCCAACTCTCCATCTGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACC 1182
DB 42646 GAGAACTCGGCAACGCGCTGCTGCTCAAGGTCAACCAAGATCGGTTCGCTGACCGAGACT 42587

QY 1183 TTGCAAGCTGTGCAATGCTCACCGGAGGGTCAACCTCACTGATGATGATGATGATGATGATG 1242
DB 42586 CTGCAAGCGCTGCAAGTGGCGGCGCAACGCGCTTCAAGTGAATGATGATGATGATGATGATG 42527

QY 1243 GGTGAGACGAGGACCAACCACTTCTGACCTCGCAGTTGCACTCACTGTGGCGCAGATC 1302
DB 42526 GCGGAGACGAGGAGCTCAACCTCGGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGATC 42467

QY 1303 AAGACTGGTGTCTCCAGCAGCTTCCGACCGTGTGCAAAAGTACAACAGGTTTCCGCGATC 1362
DB 42466 AAGACCGGCGCGCGCGCGCTCCGAGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGATC 42407

QY 1363 GAGCAGCTGTTGGGAGCGCGCGCTTACGAGGCTCGGAGGCTCGGAGGCGATTCGCGCTTTCAG 1422
DB 42406 GAGGAGATCTCGAGGACCGCGCGGTGTACGCGCGCGCGCTTCCCGCGCGCTTCCCGCTTCAAG 42347

QY 1423 GCGTAA 1428
DB 42346 GCGTGA 42341

RESULT 15
AB016822.17/c
WPCOMMENT
Sequence split into 26 fragments LOCUS AB016822 Accession AB016822
Fragment Name Begin End
AB016822_00 1 110000
AB016822_01 100001 210000
AB016822_02 200001 310000
AB016822_03 300001 410000
AB016822_04 400001 510000
AB016822_05 500001 610000
AB016822_06 600001 710000
AB016822_07 700001 810000
AB016822_08 800001 910000
AB016822_09 900001 1010000
AB016822_10 1000001 1110000
AB016822_11 1100001 1210000
AB016822_12 1200001 1310000
AB016822_13 1300001 1410000
AB016822_14 1400001 1510000
AB016822_15 1500001 1610000
AB016822_16 1600001 1710000
AB016822_17 1700001 1810000
AB016822_18 1800001 1910000
AB016822_19 1900001 2010000
AB016822_20 2000001 2110000
AB016822_21 2100001 2210000
AB016822_22 2200001 2310000

```

<p>AE016822_23 2300001 2410000 AE016822_24 2400001 2510000 AE016822_25 2500001 2584158 Continuation (18 of 26) of AE016822 from base 1700001 (AE016822 Leifsonia xyli subsp. xy)</p>		<p>Query Match 44.4%; Score 700; DB 1; Length 1100000; Best Local Similarity 71.9%; Pred. No. 9.6e-123; Matches 929; Conservative 0; Mismatches 360; Indels 3; Gaps 1;</p>	
QY	140	AGGAGGCCACAGTGGCTGAAATCATGCAGGTATTCGCTCGCGAAATTCGCACTCCCGGG	199
DB	83109	AGGAGACTGTTGGTGTGCATCGAAGCTGTAGGAGCTCGGAGATCTCTGACTCGGGG	83050
QY	200	GTAACCCAAACGTCGAGGAGAGGTTTCTCGATGACGGTTCCCAACGGTTCGCGAGTG	259
DB	83049	GAACCCGACCGTCGAGGTCGAGGTCTCTCGAAGATGGCAACGGTCAGCCGCGCGCG	82990
QY	260	TTCATCCCGCGCATCCACGGGGTCCAGAGGCTCATGAGTGGGTGAGGTGGC---G	316
DB	82989	TCCCGTCCGGCGGTCCACGGCGGTTCGAGGCTTACGAGCTCCGCGACGGGCAAGG	82930
QY	317	ATCGCTACCTGGGCAAGGGCGTTTGAAGGCAAGTTCGAAACGTCGAAAGAAATCGGGG	376
DB	82929	GCGCTACTTGGCAAGGGTTCGAGAGGCGGTGATGCCGTCTCGAGAGATCGGCC	82870
QY	377	ACGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGC	436
DB	82869	CGGCCATCGAGGGCTTCGAGGGCCAGCGACCGCGCATCGTCGATGAGGCGATGATCGAAC	82810
QY	437	TTGATGGCAACCGCAACAGTCCCGCTGGGTGCAAAACGCAATCTCTGGTGTTCGATGG	496
DB	82809	TCGACGGCACCGCAACAAAGAGCGCTGGGGCTTAACGCCATGCTCGGTGTGAGCCTCG	82750
QY	497	CTGTTGCAAGGCTGCTGTGATTCGCGAGGCTCCACCTGTTCCGCTACATCGGTGGAC	556
DB	82749	CAGTCGGAAGGCGCGCGGACCTCCGCGACCTGCGCGACTGCTGCTACCTTGGGCGGCC	82690
QY	557	CAAGCGCACAGTCTTCCAGTTTCCATGATGAACATCATCAACGGTGGCGCTCACGGTG	616
DB	82689	CGAACGCGCATGTCTGCTCCCGTCCCGATGATGAACATCATCAACGGCGCGCGACGGG	82630
QY	617	ACTCGGCTTTGAGGTTTCAAGATTCGATCGCTCCAAATCGGTGAGAGACCTTCTCTG	676
DB	82629	ACACCGGCTCGACATCCAGGATTCATGATTCCTCCGATCGCGCGCGGAGACCTTCTCG	82570
QY	677	AGGCTCTCGCAACGGCGGGAGTCTACCGCACTGAAGTCCGTCATCAAGGAAGG	736
DB	82569	AGGGCTTCGCGCTGGGGCGTGGAGACCTACCACTCGCTCAAGGCGCTGCTGAAGT	82510
QY	737	GCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCGTCGGCTCCACCCGTG	796
DB	82509	GCTTGAACACCGGCTGGGGGACGAGGGCGGCTTCGCCCCGAGCTGGAGGACACCGCG	82450
QY	797	AGGCTCTGACCTTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCCGAGGCAAGGACA	856
DB	82449	CTGGCTCGACCTGATCGCGAGGCCATCGAAGGGCGGCTTCAACGTCGGCTCGAGA	82390
QY	857	TGCTCTTGTCTGAGAGGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTACCACTCG	916
DB	82389	TGCGCTTCGGCTTCGAGTGGGCTCCACCGAGTTCTTCGAGAACGGCGCTTACCGCTCG	82330
QY	917	AAGGTGGCGAGCACTCCGCGAGCTGAGATGCGCAACGTTTACGCTGAGCTGTTGACGGT	976
DB	82329	AGGGGAGGAGACCGCACCGCGGCGAGATGAGCGCTTACTACCGATCTCGGAACT	82270
QY	977	ACCAATCGTCTCCATCGAGGACCCACTGCGAGGAGATGACTGGGAGGTTTACCAACCC	1036
DB	82269	ACCGCTGTGTTCTCATCGAGGACCCGCTGCGGAGGACGACTGGGAGGCTGGCGCACCC	82210
QY	1037	TACCGCAACCATCGGCGGACAGGTTTCAGATCGTTGGCGAGCACTTCTTCGTACCAACC	1096
DB	82209	TGAACCGCGAGATCGGCTCGAGCTGCAACTCGTCGGCGGATGACCTGTTCTCGTCAACACC	82150

Search completed: September 28, 2005, 20:45:27
 Job time : 6714 secs

QY	1097	CTGAGCGCTGAAGAGGGGCAATCGCTAAGAAAGGCTGCCAACTCCATCTCTGTTAAGGTGA	1156
DB	82149	CGAAGCGCTGGCGCAGGGCATCCAGCAGAAAGGCGCGCAACAGCATCTCTGGTCAAGGTGA	82090
QY	1157	ACCAGATCGGTACCTCTACCGAGACCTTCGAGCGTGTGACATGGCTCACCGCGCAGGCT	1216
DB	82089	ACCAGATCGGCACCTTGAACCGAGAGCTGGAGCGCGGTCTCGCTCCGCCAGCGCACGGTA	82030
QY	1217	ACACCTTCATGATGTCCCAACCGTTCCGGTGAGACCGAGGACACCACTTGTGTGACCTCG	1276
DB	82029	TGACCGCGCTCTCTCGCACCGCTCCGCGAGACCGAGGACACCACTTGTGTGACCTCG	81970
QY	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCACTTCCGACCGTGTG	1336
DB	81969	CCGTGCGCACCGATTCGCGCCAGATCAAGACCGGCGCCCGCGCTCCGAGCGCGTGG	81910
QY	1337	CAAGTACAACACGCTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTCTACGCGAG	1396
DB	81909	CGAAGTACAACGAGCTGCTCCGCACTCGAGGAGGAGCTGGGCGAGGCGCTGTCTACGCGG	81850
QY	1397	GTCGCGCGCATTCGCCACGCTTTTCAGGGCTAA	1428
DB	81849	GCGTTGCGCGCTTCCCGGTTACCGAGGCTGA	81818

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 23:24:17 ; Search time 93.5 Seconds
(without alignments)
1758.002 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VAEIMHVAREILDGRGNT.....QLLGDAVYAGSAPPRFQG 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151	99.8	425	4 AAB70882	Aab70882 C. glutam
2	2148	99.7	425	4 AAG90831	Aag90831 C. glutami
3	2148	99.7	425	6 ABP97001	Abp97001 Coryneb
4	2145	99.5	425	4 AAB79278	Aab79278 Coryneb
5	2145	99.5	425	7 ADD13323	Add13323 C. glutam
6	2144	99.5	425	6 ABP97002	Abp97002 Modified
7	1913	88.8	425	6 ABU25984	Abu25984 Protein e
8	1562	72.5	423	8 ADN25687	Adn25687 Bacterial
9	1560.5	72.4	429	6 ABU34209	Abu34209 Protein e
10	1553.5	72.1	429	6 ABU36566	Abu36566 Protein e
11	1553.5	72.1	439	6 ABU34772	Abu34772 Protein e
12	1505.5	69.9	455	4 AAU54498	Aau54498 Propionib
13	1505.5	69.9	455	6 ABM51017	Abm51017 Propionib
14	1502.5	69.7	447	6 ABU35758	Abu35758 Protein e
15	1422	66.0	423	8 ADS22663	Ads22663 Bacterial
16	1412.5	65.5	426	8 ADS27878	Ads27878 Bacterial
17	1408	65.3	429	6 ABP58571	Abp58571 Thermoana
18	1402.5	65.1	423	8 ADS21524	Ads21524 Bacterial
19	1402.5	65.1	426	8 ADS28853	Ads28853 Bacterial
20	1402	65.1	429	8 ADS28400	Ads28400 Bacterial
21	1396	64.8	424	8 ADS22786	Ads22786 Bacterial
22	1394	64.7	420	8 ADS25706	Ads25706 Bacterial
23	1394	64.7	423	8 ADS26189	Ads26189 Bacterial
24	1392.5	64.6	431	6 ABU24836	Abu24836 Protein e
25	1377.5	63.9	430	8 ADO59753	Ado59753 B. subtil

ALIGNMENTS

RESULT 1

AAB70882
ID AAB70882 standard; protein; 425 AA.

XX AC AAB70882;

XX 12-JUL-2001 (first entry)

XX C. glutamicum enolase protein.

XX Enolase; fermentation; L-amino acid; L-lysine; coryneform; eno gene;
KW medicine; animal feed supplement.

XX Corynebacterium glutamicum.

XX EPI090998-A1.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121158.

XX 05-OCT-1999; 99DE-01047791.

XX (DEGS) DEGUSSA-HUELS AG.

XX Moeckel B, Pfeifferle W, Hermann T, Puehler A, Kalinowski J;
XX Bathe B;

XX WPI; 2001-292928/31.

XX N-PSDB; AAF61696.

XX New enolase gene from coryneform bacteria, used to prepare transformants
XX with increased synthesis of amino acids, particularly lysine.

XX Claim 6; Page 14-16; 25pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria which is used in a method for fermentative production
XX of L-amino acids, especially L-lysine, by fermenting a lysine-producing
XX coryneform in which the eno gene has been amplified, and isolating amino
XX acids that have accumulated in the medium or cells. (I), which encodes an
XX enolase, is used to transform coryneforms for production of L-amino
XX acids, specifically lysine which is used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Overexpression of (I) improves production of
XX amino acids, especially of L-lysine. This sequence represents the enolase
XX protein described in the method of the invention

AdS44869 Bacterial
Abu18243 Protein e
Abu25001 Protein e
AdS43300 Bacterial
AdS42576 Bacterial
Abu48815 Listeria
Abu33030 Protein e
AdS30458 Bacterial
Abu23996 Protein e
Abu19880 Protein e
Adn26444 Bacterial
Adn25109 Bacterial
Adn22035 Bacterial
Abu21358 Protein e
Adn24794 Bacterial
Adr44418 M. methyl
Aau33516 Enterococ
Aau35261 Enterococ
Abu14630 Protein e
Adn17379 Bacterial

26 1377.5 63.9 430 8 ADS44869
27 1373 63.7 431 6 ABU18243
28 1366.5 63.4 448 6 ABU25001
29 1362.5 63.2 426 8 ADS43300
30 1362.5 63.2 427 8 ADS42576
31 1359.5 63.1 430 5 ABU48815
32 1359.5 63.1 430 6 ABU33030
33 1357.5 63.0 430 8 ADS30458
34 1354.5 62.9 431 6 ABU23996
35 1351.5 62.7 427 6 ABU19880
36 1347.5 62.5 417 8 ADN26444
37 1347.5 62.5 424 8 ADN25109
38 1347.5 62.5 426 6 ADN22035
39 1347.5 62.5 427 6 ABU21358
40 1347.5 62.5 427 8 ADN24794
41 1342.5 62.3 428 8 ADR44418
42 1342 62.3 429 4 AAU33516
43 1342 62.3 432 4 AAU35261
44 1342 62.3 432 6 ABU14630
45 1334.5 61.9 426 8 ADN17379

SQ Sequence 425 AA;
 Query Match 99.8%; Score 2151; DB 4; Length 425;
 Best Local Similarity 99.8%; Pred. No. 2.6e-169;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLG 60
 QY 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRILGANAILGVSMVAKA 120
 DB 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRILGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDQVQEMPIAGETFSALRN 180
 DB 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDQVQEMPIAGETFSALRN 180
 QY 181 GAENVHAKSVIKEGLSTGLDGGFPAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
 DB 181 GAENVHAKSVIKEGLSTGLDGGFPAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
 QY 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSMM 360
 DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPASDRVAKYQNLRLRIEQLLDGAGVYAGRSF 420
 DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPASDRVAKYQNLRLRIEQLLDGAGVYAGRSF 420
 QY 421 PRFQ 425
 DB 421 PRFQ 425
 RESULT 2
 AAG90831 standard; protein; 425 AA.
 ID AAG90831
 AC AAG90831;
 DT 26-SEP-2001 (first entry)
 DE C glutamicum protein fragment SEQ ID NO: 4585.
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 OS Corynebacterium glutamicum.
 FN EP1108790-A2.
 PD 20-JUN-2001.
 PF 18-DEC-2000; 2000EP-00127688.
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000TP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI: 2001-376931/40.
 XX N-PSDB; AAH66050.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 PS Claim 17; SEQ ID NO 4585; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX Sequence 425 AA;
 SQ Query Match 99.7%; Score 2148; DB 4; Length 425;
 Best Local Similarity 99.5%; Pred. No. 4.5e-169;
 Matches 423; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLG 60
 DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLG 60
 QY 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRILGANAILGVSMVAKA 120
 DB 61 KGVLKAVENNEEIGDELADQRLIDEAMIKLDTANKSRILGANAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDQVQEMPIAGETFSALRN 180
 DB 121 AADSAGLPLFRYIGGPNHVLVPMNIIINGGAHDSGVVDQVQEMPIAGETFSALRN 180
 QY 181 GAENVHAKSVIKEGLSTGLDGGFPAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
 DB 181 GAENVHAKSVIKEGLSTGLDGGFPAPSVGSTREALDLIVKAEKAGTTPGKDIALAL 240
 QY 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTATI 300
 DB 241 DVASSEFFKQGYHFEFGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTATI 300
 QY 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSMM 360
 DB 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSMM 360
 QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPASDRVAKYQNLRLRIEQLLDGAGVYAGRSF 420
 DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPASDRVAKYQNLRLRIEQLLDGAGVYAGRSF 420
 QY 421 PRFQ 425
 DB 421 PRFQ 425
 RESULT 3
 ABP97001 standard; protein; 425 AA.
 ID ABP97001
 AC ABP97001;
 XX 17-JUN-2003 (first entry)
 XX Corynebacterium glutamicum eno protein SEQ ID NO:2.
 DE Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno.
 XX Corynebacterium glutamicum.

PT . modulation of production of fine chemicals e.g. amino acids.

PT carbohydrates or enzymes.
 XX
 PS Claim 20; Page 244-245; 1246pp; English.
 XX
 CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localize C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)
 XX
 SQ Sequence 425 AA;

Query Match 99.5%; Score 2145; DB 4; Length 425;
 Best Local Similarity 99.5%; Pred. No. 8.1e-169;
 Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREIILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 DB 1 VAEIMHVFAREIILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 QY 61 KGVLKAVENNVNEIIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 DB 61 KGVLKAVENNVNEIIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 DB 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIITGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 QY 181 GAENVYHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGFTPGKDTALAL 240
 DB 181 GAENVYHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGFTPGKDTALAL 240
 QY 241 DVASSFFKDGTHFEGGQHSAAEMANVYAEVLDAVPIVSIEDPLQDDWEGVTNLTATI 300
 DB 241 DVASSFFKDGTHFEGGQHSAAEMANVYAEVLDAVPIVSIEDPLQDDWEGVTNLTATI 300
 QY 301 GDRKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360
 DB 301 GDRKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGYTSMM 360
 QY 361 SHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNNQILLRIEQLLDGAGYAGRSF 420
 DB 361 SHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNNQILLRIEQLLDGAGYAGRSF 420
 QY 421 PRFQG 425
 DB 421 PRFQG 425

RESULT 5
 ADD13323
 ID ADD13323 standard; protein; 425 AA.
 XX
 AC ADD13323;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE C. glutamicum carbon metabolism associated protein RXA00235.
 XX carbon metabolism; energy-rich molecule; oxidative phosphorylation;

KW fine chemical; amino acid production; lysine production;
 KW nucleotide production; nucleoside production; lipid production;
 KW fatty acid production; diol production; carbohydrate production;
 KW aromatic compound production; vitamin production; co-factor production;
 KW enzyme production; food; animal feed; cosmetic; pharmaceutical.
 XX
 OS Corynebacterium glutamicum.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 223 /note= "Optionally substituted with Lys"
 XX
 PN WO2003040291-A2.
 XX
 PD 15-MAY-2003.
 XX
 PP 31-OCT-2002; 2002WO-EP012135.
 XX
 PR 05-NOV-2001; 2001DE-01054270.
 XX
 PA (BADI) BASF AG.
 XX
 PI Zelder O, Pompejus M, Schroeder H, Kroeger B, Kloppeogge C;
 PI Haberhauer G;
 XX
 DR WPI; 2003-505068/47.
 DR N-PSDB; ADD13322.
 XX
 PT New nucleic acid encoding variant forms of proteins involved in carbon
 PT compound metabolism, useful for production of fine chemicals,
 PT specifically lysine, in microorganisms.
 XX
 PS Claim 1; SEQ ID NO 8; 259pp; German.
 XX
 CC This invention describes novel polynucleotides and polypeptides
 CC associated with the metabolism of carbon compounds and generation of
 CC energy-rich molecules by oxidative phosphorylation in Corynebacterium
 CC glutamicum. The polynucleotides of the invention are isolated from a
 CC nucleic acid library of C. glutamicum then mutated at the specified
 CC positions, cloned and expressed by standard methods. Cells, especially
 CC Corynebacterium glutamicum, containing vectors that express the
 CC polynucleotides are used for production of fine chemicals, preferably
 CC amino acids and specifically lysine, but more generally nucleotides,
 CC nucleosides, lipids, fatty acids, diols, carbohydrates, aromatic
 CC compounds, vitamins, co-factors and enzymes. These are useful in the
 CC food, animal feed, cosmetics and pharmaceutical industries.
 CC Polynucleotides, optionally as primers and probes, can also be used for
 CC identification and classification of C. glutamicum and related species,
 CC e.g. for diagnosis, for genomic mapping; functional or evolutionary
 CC studies gene manipulation, and modulation of metabolic activity. Cells
 CC that contain the products of the invention may produce fine chemicals in
 CC better yields, with higher productivity and/or more efficiently.
 XX
 SQ Sequence 425 AA;

Query Match 99.5%; Score 2145; DB 7; Length 425;
 Best Local Similarity 99.5%; Pred. No. 8.1e-169;
 Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREIILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 DB 1 VAEIMHVFAREIILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
 QY 61 KGVLKAVENNVNEIIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 DB 61 KGVLKAVENNVNEIIGDELADDDORLIDEAMIKLDGTANKSRGKAGNAILGVSMVAKA 120
 QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIINGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 DB 121 AADSAGLPLFRYIGGPNNAHVLVPMNNIITGGAHDSGVVDVQEFMIAPIGAEFTFSALRN 180
 QY 181 GAENVYHALKSVIKKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGFTPGKDTALAL 240

Db 181 GAETHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHPEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHPEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETPDVDMHRAGYTSM 360
 Db 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETPDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 QY 421 PRFOG 425
 Db 421 PRFOG 425

RESULT 6
 ABP97002
 ID ABP97002 standard; protein; 425 AA.
 XX
 AC ABP97002;
 DT 17-JUN-2003 (first entry)
 DE Modified eno S330E protein SEQ ID NO:6.
 XX

KW Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno.
 XX

OS Corynebacterium glutamicum.
 OS Synthetic.
 XX

FN WO2003023016-A2.
 XX

PD 20-MAR-2003.
 XX

PF 11-SEP-2002; 2002WO-EP010174.
 XX

PR 13-SEP-2001; 2001DE-01045043.
 XX

XX (DEGS) DEGUSSA AG.
 FA

XX Farwick M, Hermann T;
 FI

XX WPI; 2003-354534/33.
 DR

XX N-P5DB; ACC45314.
 XX

PT Microorganism useful for producing e.g. fine chemicals, has permanently
 PT altered phosphorylatability protein, such that biosynthesis of fine
 PT chemical synthesized by microorganism is increased compared to wild-type.
 XX

PS Example 3; Page 85-86; 120pp; English.
 XX

CC The present invention describes a microorganism (I), in which the
 CC phosphorylatability of at least one protein has been permanently altered
 CC such that the biosynthesis of at least one fine chemical synthesised by
 CC the microorganism is increased compared to the wild type. Also described:
 CC (1) use of a DNA (II) sequence coding for a protein which contains a
 CC phosphorylation site, where the sequence contains such a mutation that
 CC the protein is changed in its phosphorylatability for the production of
 CC (I), or for the production of fine chemicals; and (2) a method for
 CC producing fine chemicals or metabolites comprising using (I). (I) is
 CC useful for producing fine chemicals or metabolites, such as amino acids,
 CC vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
 CC and vitamins produced using (I) can be used in human medicine, in the
 CC pharmaceutical industry, food industry and in animal feeding. (I)
 CC produces larger amount of desired fine chemical or a metabolite than the
 CC wild type. The present sequence represents a modified eno S330E from

CC Corynebacterium glutamicum, which is used in an example from the present
 CC invention
 XX
 SQ Sequence 425 AA;
 Query Match 99.5%; Score 2144; DB 6; Length 425;
 Best Local Similarity 99.3%; Pred. No. 9.8e-169;
 Matches 422; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VAEIMHVPAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 Db 1 MAEIMHVPAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYL 60
 QY 61 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRLGANAILGVSNVAKA 120
 Db 61 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRLGANAILGVSNVAKA 120
 QY 121 AADSAGLPLFRYIGGPNHVLVPMVMNIINGGAHADSGVDVQEFMIPIGAETFEALRN 180
 Db 121 AADSAGLPLFRYIGGPNHVLVPMVMNIINGGAHADSGVDVQEFMIPIGAETFEALRN 180
 QY 181 GAEVYHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 Db 181 GAEVYHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVEAIEKAGFTPGKDIALAL 240
 QY 241 DVASSEFFKDGTYHPEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 Db 241 DVASSEFFKDGTYHPEGGQSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATI 300
 QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETPDVDMHRAGYTSM 360
 Db 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETPDVDMHRAGYTSM 360
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDAGVYAGRSAP 420
 QY 421 PRFOG 425
 Db 421 PRFOG 425

RESULT 7

ID ABU25984 standard; protein; 425 AA.

XX ABU25984;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #11511.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Corynebacterium diphtheriae.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI: 2003-029926/02.
 DR N-PSDB; ACA38842.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 62696; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 439 AA;

Query Match 72.1%; Score 1553.5; DB 6; Length 439;
 Best Local Similarity 73.2%; Pred. No. 7,9e-120;
 Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVPAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAEHRLDGGDRLGKGV 63
 DB 14 IEQVGAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAEHRLDGGDRLGKGV 73
 QY 64 LKAVENVEEIGDELAGEADDORLIDRAMIKLDTANKSRIGANVILGVSMVAKAARD 123
 DB 74 QKAVQVLEIGLGPVILGNLADDORLIDRAMIKLDTANKSRIGANVILGVSMVAKAARD 133
 QY 124 SAGLPLFRYIGPNNAHLPVPMNNINGGAHDSGVYDQEFMFIAPIGATFSEALRNGAE 183
 DB 134 SAEPLFLFRYIGPNNAHLPVPMNNINGGAHDSGVYDQEFMFIAPIGATFSEALRNGAE 193
 QY 184 VTHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGFTFGKIDIALDVA 243
 DB 194 VTHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGFTFGKIDIALDVA 253
 QY 244 SSEFFKCT-VHFEQGHSAEAMVYAEVLDVAYPIVSLDPLQDDWEGVYTNLTATGD 302
 DB 254 ATEFFTDGTGVYFEGTADQMTBFFYAGLGLAYPLVSDIEDPLSDDDWGAALTAISGD 313
 QY 303 KVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSMHS 362
 DB 314 RVQIVGDDIFVTNPERLKEGIERGVANALLVKVNOIGTLTETLDVTLAHGGVYRTMISH 373

QY 363 RSGTEDTTIADLAVLNCGOIKTGAPARSDRVAKYNOLLRIEQLGDAVYAGSAPPR 422
 DB 374 RSGTEDTTIADLAVLNCGOIKTGAPARSDRVAKYNOLLRIEQLGDAVYAGSAPPR 433
 QY 423 F 423
 DB 434 F 434
 RESULT 12
 AAU54498
 ID AAU54498 standard; protein; 455 AA.
 AC AAU54498;
 XX
 DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #15394.
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 OS Propionibacterium acnes.
 XX WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PP 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonmeuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59564.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Claim 6; SEQ ID NO 15693; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 455 AA;

PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 11696; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 423 AA;

Query Match	66.0%;	Score 1422;	DB 8;	Length 423;
Best Local Similarity	66.2%;	Pred. No. 5.8e-109;		
Matches 276;	Conservative 58;	Mismatches 81;	Indels 2;	Gaps 1;
Qy	4	IMHVFARILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGGDRYLKGV	63	
Db	4	ITDIAREILDSRGNPTVEVDVYLEDGSMGRAAVPSGASTGAHEAVELRDGKRYLGKV	63	
Qy	64	LKAVENVEITGDELADQRLIDEAMTKLDGTANKSLGANAILGVSMVAKAAAD	123	
Db	64	EKAVEAVNTEIFDAIGDFDAENQIQIDQMIALDGTNKSRLGANAILGVSLAKAAAE	123	
Qy	124	SAGLPLFRYIGGPNNAHVLVPMNIIINGGAHADSGVDVQEFMIAPIGATFSEALRNGAE	183	
Db	124	ASGLPLRYRYVGGPNNAHLLVPMNIIINGGAHADNPIDQEFMILPVGAENIRRAVRMGSE	183	
Qy	184	VYHALKSVIKEGLSTGLDEGGFAPSVGSTREALDLIVKAEKAGFTPGKDIALALDVA	243	
Db	184	VFHTLKKELSAQQHNTNVGDEGGFAPGLSSAPALDFIMKSIKAGYRPGEDMYVGLDCA	243	
Qy	244	SSPFFKDGTYHFEGGQHS--AAEMANVYAEIVDAPIVTSIEDPLQEDDEGYTNLTATIG	301	
Db	244	STFFKDGKTVLEGGEGTLEPGAWAYLAELVNKYPISVEDGHAEDDEGWTLTDLVG	303	
Qy	302	DKQIVGDDPFVTNPERLKEGIAKKAANSILVKVNQIGTTLTETPDVDMAHRAGYTSMMS	361	
Db	304	NKCQLVGDDLPVTNSARLRDGIKMGVANSILVKVNQIGSLSETLDAVETAHKAGYTAVMS	363	
Qy	362	HRSGETEDTTIADLAVNLNGQIKTGAPASDRVAKNQLIRIQLLDGAGVYAGRS	418	
Db	364	HRSGETEDSTIADLAVATNCQINTGSLARSRLAKYNQLIRIEEMILGPOAYAGRS	420	

Search completed: September 29, 2005, 12:10:39
Job time : 94.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:08:34 ; Search time 275.5 Seconds
(without alignments)
9372.213 Million cell updates/sec

Title: US-10-728-947-1

Perfect score: 1578

Sequence: 1 99ctg99gatgggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	4	US-09-860-768-1
2	1576.4	99.9	1578	4	US-09-860-768-3
3	632.6	40.1	4403765	3	US-09-103-840A-2
4	632.6	40.1	4411529	3	US-09-103-840A-1
5	552	35.0	1299	4	US-09-902-540-4780
6	552	35.0	27903	4	US-09-902-540-1235
7	533.6	33.8	1335	4	US-09-252-991A-4449
8	533	33.8	1347	4	US-09-252-991A-4158
9	481.2	30.5	1380	4	US-09-489-039A-2890
10	449.2	28.5	2363	4	US-09-818-780-22
11	442.4	28.0	3545	2	US-09-781-802-9
12	433.2	27.5	1305	4	US-09-583-110-1156
13	432	27.4	11384	3	US-08-961-527-45
14	429.6	27.2	1196	4	US-09-107-532A-156
15	420.2	26.6	1233	4	US-09-489-039A-2858
16	415.6	26.3	1511	3	US-08-847-065-18
17	415.6	26.3	1511	4	US-09-829-382-18
18	391.8	24.8	3513	3	US-09-058-260-1
19	391.8	24.8	4315	2	US-08-781-802-1
20	391.8	24.8	4315	3	US-08-694-078-1
21	391.2	24.8	1830121	4	US-09-557-884-1
22	391.2	24.8	1830121	4	US-09-543-990A-1
23	395.4	24.4	1371	4	US-09-543-681A-1918
24	393.2	24.3	1224	4	US-09-107-433-2383
25	374.8	23.8	1335	4	US-09-328-352-3143
26	367.2	23.3	1338	4	US-09-540-236-434
27	367.2	23.3	92407	4	US-09-596-002-36

28	359.6	22.8	1460	4	US-09-830-217-21	Sequence 21, Appl
29	359.6	22.8	1460	4	US-10-278-946-21	Sequence 21, Appl
30	349.6	22.2	1305	4	US-09-710-279-2321	Sequence 221, Ap
31	349.6	22.2	1323	3	US-09-134-001C-2804	Sequence 2804, Ap
32	349.6	22.2	3830	4	US-09-710-279-3773	Sequence 3773, Ap
33	344.6	21.8	1539	4	US-08-956-171E-207	Sequence 207, App
34	344.6	21.8	1539	4	US-08-781-986A-207	Sequence 207, App
35	340.6	21.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
36	330.6	21.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
37	330.6	21.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
38	330.2	20.9	1230025	4	US-09-198-452A-1	Sequence 1, Appli
39	309.6	19.6	1314	4	US-09-614-221A-509	Sequence 509, App
40	306.4	19.4	1314	4	US-09-614-221A-466	Sequence 466, App
41	303.2	19.2	1884	4	US-09-566-921-38	Sequence 38, Appl
42	302.6	19.2	934	3	US-08-961-083-183	Sequence 183, App
43	302.6	19.2	934	4	US-09-536-784-183	Sequence 183, App
44	301.2	19.1	2422	4	US-09-949-016-282	Sequence 282, App
45	301.2	19.1	2422	4	US-09-949-016-2488	Sequence 2488, Ap

ALIGNMENTS

RESULT 1
US-09-860-768-1
; Sequence 1, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860.768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-1

Query Match	100.0%	Score 1578;	DB 4;	Length 1578;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1578;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATTGCTCATCGAACAAAGA	60	
Db	1	GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATTGCTCATCGAACAAAGA	60	
Qy	61	TTCTGTGCAACAATTTGGGTAGAGCTGATTGAAGACATTTTGCATCGTGAATAATTTCTAG	120	
Db	61	TTCTGTGCAACAATTTGGGTAGAGCTGATTGAAGACATTTTGCATCGTGAATAATTTCTAG	120	
Qy	121	TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGTGCTGAATATCATGCGTATTCGTCGC	180	
Db	121	TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGTGCTGAATATCATGCGTATTCGTCGC	180	
Qy	181	GAATTTCTCCACTCCCGCGGTACCCCAACCGTTCGAGCAGAGGTTTTCCTGGATGACGGT	240	
Db	181	GAATTTCTCCACTCCCGCGGTACCCCAACCGTTCGAGCAGAGGTTTTCCTGGATGACGGT	240	
Qy	241	TCCACAGCTGTCGAGGTGTTTCATCCCGCGCATCCACCGCGCTCCACAGGCTCATGAG	300	
Db	241	TCCACAGCTGTCGAGGTGTTTCATCCCGCGCATCCACCGCGCTCCACAGGCTCATGAG	300	
Qy	301	CTGCGTACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGAGGAGTTGAAAACGTC	360	
Db	301	CTGCGTACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGAGGAGTTGAAAACGTC	360	

```
QY 361 AACGAAGAAATCGGCGAGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAAGAAATCGGCGAGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGTGATGGCAGCGCAAAAGTCCCGCTCGGCTGCAAAAGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGTGATGGCAGCGCAAAAGTCCCGCTCGGCTGCAAAAGCAATC 480
QY 481 CTTGGTGTTCCTTCCATGGCTGTGTCAGAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTT 540
Db 481 CTTGGTGTTCCTTCCATGGCTGTGTCAGAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTT 540
QY 541 CGCTACATCGGTGACCAAAAGCGACACAGTTCCTCCAGTTCCATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGACCAAAAGCGACACAGTTCCTCCAGTTCCATGATGAACATCATCAAC 600
QY 601 GGTGGCGCTCAGCTGACTCGGCTGTTGAGCTTCAGGAAATTCATGATCGCTCCAATCGGT 660
Db 601 GGTGGCGCTCAGCTGACTCGGCTGTTGAGCTTCAGGAAATTCATGATCGCTCCAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGCTTACCAACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGCTTACCAACGCACTGAAGTCC 720
QY 721 GTCATCAAGGAAAGGCGCTGTCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGGAAAGGCGCTGTCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
QY 781 GTGGGCTCCACCGTGAAGGCTCTTGCATTCATGAGGCAATCGAAGAGGCTGGGCTTC 840
Db 781 GTGGGCTCCACCGTGAAGGCTCTTGCATTCATGAGGCAATCGAAGAGGCTGGGCTTC 840
QY 841 ACCCGAGCAAGGACATCGCTCTTGTCTGAGAGTGTCTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCGAGCAAGGACATCGCTCTTGTCTGAGAGTGTCTCTCTGAGTTCTTCAAGGAC 900
QY 901 GGCACCTACCACTTCGAAGTGGCGAGCATCCCGAGCTGAGATGGCAAGGTTTACGCT 960
Db 901 GGCACCTACCACTTCGAAGTGGCGAGCATCCCGAGCTGAGATGGCAAGGTTTACGCT 960
QY 961 GAGCTCGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGTACCCCAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGACTGG 1020
QY 1021 GAGGTTTACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAGCAC 1080
Db 1021 GAGGTTTACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGAGCAC 1080
QY 1081 TTCTTCGTCAACCAACCTGAGCGCTGAGGAGGGCATCGCTAAGAGGCTGCCAATCC 1140
Db 1081 TTCTTCGTCAACCAACCTGAGCGCTGAGGAGGGCATCGCTAAGAGGCTGCCAATCC 1140
QY 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTCCAGAGCTTCGAGCGCTGTGCAATG 1200
Db 1141 ATCTCGTTAAGGTGAACAGATCGGTACCTCCAGAGCTTCGAGCGCTGTGCAATG 1200
QY 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 1260
Db 1201 GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACACC 1260
QY 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGCTGTGCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGCTGTGCTCCAGCA 1320
QY 1321 GGTTCGAGCCGTGTGCGAAAGTACAAACAGGCTTCTCCGCACTCGAGCAGCTGCTGGCGAC 1380
Db 1321 GGTTCGAGCCGTGTGCGAAAGTACAAACAGGCTTCTCCGCACTCGAGCAGCTGCTGGCGAC 1380
QY 1381 GCCGCGCTTACGAGGTGCGAGCGCATTCGCCACGCTTTTACGGGCTAAATAAAGCGCTTT 1440
Db 1381 GCCGCGCTTACGAGGTGCGAGCGCATTCGCCACGCTTTTACGGGCTAAATAAAGCGCTTT 1440
```

```
QY 1441 TTCGAGCCCGGTAACTCAAGTTGCCGGGCGCTCGTTGCTTACTACTGTACTGTGT 1500
Db 1441 TTCGAGCCCGGTAACTCAAGTTGCCGGGCGCTCGTTGCTTACTACTGTACTGTGT 1500
QY 1501 GACTATGATCGAGGATTATGGCAAGACGAAAGAAAACCTATAAAGGCTTGTTCCTGTCT 1560
Db 1501 GACTATGATCGAGGATTATGGCAAGACGAAAGAAAACCTATAAAGGCTTGTTCCTGTCT 1560
QY 1561 CAAGCAGGAAACGTGCTT 1578
Db 1561 CAAGCAGGAAACGTGCTT 1578

RESULT 2
US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match 99.9%; Score 1576.4; DB 4; Length 1578;
Best Local Similarity 99.9%; Pred.No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
Db 1 GGCTGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
QY 61 TTCTGTCACAAATTTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATATTTCTAG 120
Db 61 TTCTGTCACAAATTTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATATTTCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACTATTCGCTCGC 180
QY 181 GAAATTCCTGACTCCCGCGGTAAACCCAAACCGTCGAGGACAGGTTTTCCTGGATGACGGT 240
Db 181 GAAATTCCTGACTCCCGCGGTAAACCCAAACCGTCGAGGACAGGTTTTCCTGGATGACGGT 240
QY 241 TCCACGAGTTCGAGAGTGTTCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300
Db 241 TCCACGAGTTCGAGAGTGTTCATCCGGCGCATCCACCGGCTCCACGAGGCTCATGAG 300
QY 301 CTGCGTCACGCTGGCGCATCGCTACCTGGCAAGGGCGTCTTTGAAGGAGTGTGAAAACGTC 360
Db 301 CTGCGTCACGCTGGCGCATCGCTACCTGGCAAGGGCGTCTTTGAAGGAGTGTGAAAACGTC 360
QY 361 AACGAAGAAATCGGCGAGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAAGAAATCGGCGAGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGTGATGGCAGCGCAAAAGTCCCGCTCGGCTGCAAAAGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGTGATGGCAGCGCAAAAGTCCCGCTCGGCTGCAAAAGCAATC 480
QY 481 CTTGGTGTTCCTTCCATGGCTGTGTCAGAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTT 540
```

Db 481 CTTGGTGTTCATGGCTGTTGCAAGAGCTCTGCTGATTCCGAGGCTCCCACTGTTC 540
Qy 541 CGCTACATCGGTGACCAAGCGCACACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGACCAAGCGCACACGTTCTTCCAGTTTCCAATGATGAACATCATCAAC 600
Qy 601 GGTGGCGCTACGCTGACTCCGGTGTGACGTTTCCAGGAATTCATGATCGCTCCAAATCGGT 660
Db 601 GGTGGCGCTACGCTGACTCCGGTGTGACGTTTCCAGGAATTCATGATCGCTCCAAATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCGGCAACCGGCGGAGGTCTACACGACCTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCGGCAACCGGCGGAGGTCTACACGACCTGAAGTCC 720
Qy 721 GTCATCAAGGAAAGGGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCCTTCC 780
Db 721 GTCATCAAGGAAAGGGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCCTTCC 780
Qy 781 GTGGGCTCCACCGGTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC 840
Db 781 GTGGGCTCCACCGGTGAGGCTCTTGAACCTTATCGTTAAGGCAATCGAAGGCTGGCTTC 840
Qy 841 ACCCGAGCAAGGACATCGCTCTTGTCTGCAAGTTCCTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCGAGCAAGGACATCGCTCTTGTCTGCAAGTTCCTCTCTGAGTTCTTCAAGGAC 900
Qy 901 GGCACTTACCACTTCCGAAGGTGGCGACGACCTCCGACGCTGAGATGGCAAGGTTTACGCT 960
Db 901 GGCACTTACCACTTCCGAAGGTGGCGACGACCTCCGACGCTGAGATGGCAAGGTTTACGCT 960
Qy 961 GAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGGTACCCAAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
Qy 1021 GAGGTTTACCAACCTTACCGGCAACCATCGGCGACAAGGTTTCAAGTTCGTTGGCGAC 1080
Db 1021 GAGGTTTACCAACCTTACCGGCAACCATCGGCGACAAGGTTTCAAGTTCGTTGGCGAC 1080
Qy 1081 TTCTTGTGTCAACACCTTACGGCTGAGGCGCTGAAGGAGGCGATCGTTAAGAGGCTGCCAATCC 1140
Db 1081 TTCTTGTGTCAACACCTTACGGCTGAGGCGCTGAAGGAGGCGATCGTTAAGAGGCTGCCAATCC 1140
Qy 1141 ATCTGTGTTAAGGTGAACAGATCGGTACCTTCAACGAGACTTTCGACGCTGTGCAATG 1200
Db 1141 ATCTGTGTTAAGGTGAACAGATCGGTACCTTCAACGAGACTTTCGACGCTGTGCAATG 1200
Qy 1201 GCTCACCGCGAGGCTACCTCCATGATGTCCCAACCGTTCCGTTGAGACCGAGGACACC 1260
Db 1201 GCTCACCGCGAGGCTACCTCCATGATGTCCCAACCGTTCCGTTGAGACCGAGGACACC 1260
Qy 1261 ACCATTGCTGACCTCGAGTTGCACTCACTGTTGGCCAGATCAAGACTGGTCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGAGTTGCACTCACTGTTGGCCAGATCAAGACTGGTCTCCAGCA 1320
Qy 1321 CGTTCCGACCGTGTGCGAAGGTACAACAGGCTTCTCCGCACTCGAGGAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGTGTGCGAAGGTACAACAGGCTTCTCCGCACTCGAGGAGCTGCTTGGCGAC 1380
Qy 1381 GCCGGGCTTACGAGGTCGAGGCGAATCCCAACGCTTTTCAAGGCTAATAAAGGCGCTT 1440
Db 1381 GCCGGGCTTACGAGGTCGAGGCGAATCCCAACGCTTTTCAAGGCTAATAAAGGCGCTT 1440
Qy 1441 TTCCAGCCCGGTAACTCAAGTTTGGCGGCGTGTGGCTTACTACTGTACTGTGT 1500
Db 1441 TTCCAGCCCGGTAACTCAAGTTTGGCGGCGTGTGGCTTACTACTGTACTGTGT 1500
Qy 1501 GACTATGATCAGGATTTATGCAAGGAGGAGGAACTCATAAAGGCGCTTCTCTGTCT 1560
Db 1501 GACTATGATCAGGATTTATGCAAGGAGGAGGAACTCATAAAGGCGCTTCTCTGTCT 1560
Qy 1561 CAAGCAGGGAACGTGCTT 1578
|||||

Db 1561 CAAGCAGGGAACGTGCTT 1578
RESULT 3
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Query Match 40.1%; Score 632.6; DB 3; Length 4403765;
Best Local Similarity 68.4%; Pred. No. 9.4e-166;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;
Qy 140 AGGAGGCCACAGTGGCTGAAATCATGCACGATTTCGTCGCGAAATTCGACCTCCCGG 199
Db 1144576 AGGAGAACCCAGTCCCGATTATCGAGCAGGTTAGGGCCCGAGAGATCCTGATTCCCGCG 1144635
Qy 200 GTAACCCAAACCGTCGAGCGAGAGGTTTCTCTGAGATGACGGTTCCACCGTGTCCGAGGTG 259
Db 1144636 GCAACCCGACGGTGGAGTTCGAGGTGGCGCTTATCGAGGGACATTCGCCCGGCCCGGG 1144695
Qy 260 TTCCATCCGCGCATCCACCGGCTCCACGAGGCTCATGAGCTCGGTGACCGTGGCGATC 319
Db 1144696 TGCGTCCGCGCTCCGACCGGGGAGCACGAGGCCGTCGAGTTGCGCACGCGCGCGATC 1144755
Qy 320 GCTACCTGGGCAAGGGCGTTTGAAGCAGTTTGAAGAGTTGAAAGCTCAACGAAGAAATCGCGACG 379
Db 1144756 GCTACGGCGCAAGGGCGTCCAAAAGCCGTCGAGGCTGTTCTTGTAGATGAGATCGGCCCGG 1144815
Qy 380 AGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGACGAAGCAATGATCAAGCTTG 439
Db 1144816 CCGTCATCGAATCAACGCCGACGACGACGATTTGGTCGACGAGCGCTGGTGACCTAG 1144875
Qy 440 ATGSCACCGCCAAACAACTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATTTGGCTG 499
Db 1144876 ACGGCACCCCGCAAGTCCCGGCTGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 1144935
Qy 500 TTGCAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCCGCTGGAACCA 559
Db 1144936 TTGCAAGGCGGCGCGGATTCGCGGAGCTGCGGTTGTTCCGTTATGTCTCGGGGGGCCAA 1144995
Qy 560 ACGCACACGTTCTTCCAGTTTCAATGATGAACATCATCAACGGTGGGCTCACGCTGACT 619
Db 1144996 ACGGGACATCTTCCCGGTACCGATGATGAACATCTCAACGGCGGCGACACGCCGATA 1145055
Qy 620 CCGGTGTGAGTTTCAAGAAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGG 679
Db 1145056 CCGGTGTGCAATTCAGAGTTTCAATGTGGCGCAATTGGCGCGCCAGCTTCGTCGAGG 1145115
Qy 680 CTCTCCCAACGGCGCGAGGTCTTACCGCATGAAAGTCCGTATCAAGGAAAGGGCC 739
Db 1145116 CGTTGCGCTGGGCGCTGAGGTGTACCAACGCGCTCAAGTCGCTCTCAAGAAAGAGGGCGC 1145175
|||||

740 TGTCCACGGAGCTTGGGATGAGGCGCTTCGCTCCTTCGTCGGCTCCACCGTGGAG 799
 1145176 TGTCCACGGGCTGGGAGCGAAGCGGCTTCGCCCGGATGTGCGCGGACCAACCGCG 1145235
 800 CTCTTGACCTTATCGTTGAGCAATCGAGAAGGCTTGCTTACCCCAAGCAAGACATCG 859
 1145236 CGTTGGACCTGATCAGCCGGGCGATCGAGTCGGCGGGCTTCGCGACCCCGGCGCGAGCTGG 1145295
 860 CTCTTGCTCTGGAGCTTGTCTCTCTGAGTCTTCAAGGACGGAC---CTACCACTTCG 916
 1145296 CGTGGCCCTTGAGCGCGCGGCGCCACCGAGTCTTTCACCGACGGACCGGCTACGCTTCG 1145355
 917 AAGTGGCCAGCACTCCGACGTCAGATGGCAACGTTTACGCTGAGCTGTTGACCGCT 976
 1145356 AGGCAACACCCGTACCGCAGACAGATGACCGAGTCTACCGCGGCTGCTCGCGGCT 1145415
 977 ACCCAATCGTCTCATCGAGACCCACTGCAAGGAAGATGACTGGAGGGTTTACACCAACC 1036
 1145416 ACCCGCTGGTGTGATCGAAGACCCACTGTCCGAAGACGATTGGGACGGCTGGGCGCGC 1145475
 1037 TCACCGCAACCATCGGCGACAGGTTTTCAGATCGTTGGGACGACTTCTTGTGACCAACC 1096
 1145476 TGACGGCTCGATCGGTGACCGGGTGCAATCGTCGGCGACGACATCTTTGTCAACCAATC 1145535
 1097 CTGAGCGCTGAAGGAGGCGATCGCTAAGAAAGGCTGCCAATCCATCTGTTAAAGTGA 1156
 1145536 CCGAGCGCTCGAGAGGGGATCGAACGGGCGTGGCAATGCTGCTGGTCAAGGTGA 1145595
 1157 ACCAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTGACATGGCTTACCGGCGAGGCT 1216
 1145596 ACCAGATCGGAGCGTTGACCGAGACACTTCGACGCGGTCAAGCTGGCTCACACGGCGAT 1145655
 1217 ACACCTCCATGATGTCACCGCTTCGCTGAGAGCGGAGGACACCACTTGTGACCTCG 1276
 1145656 ACCGACGATGATCAGTCAACCGAGTGGCGAGAGCGGAGACACCATGATCGCGACCTCG 1145715
 1277 CAGTTGCACTCAACTGTGGCAGATCAAGACTGTGTCTCCAGACAGTTTCGACCGTGTG 1336
 1145716 CGGTGGCATCGGAGCGGCGAGATCAAGACGGGCGGCTGCTCGCAGTGAGCGCTCG 1145775
 1337 CAAGTACACCAAGCTTCTCGCATCGAGACGCTGTGCGAGCGCGGCGTCTACGCGAG 1396
 1145776 CAATAACACCAAGCTGCTCGGATCGAAGAGGCGCTTGGCGAGCGGCGCGCTACGCGG 1145835
 1397 GTCCAGCGGATTCACCGCTTTCAGGCTTAATAAAGCGCT 1439
 1145836 GCGACCTGGCATTTCTCCGTTTCGCTGCGAGACGAAATAGGT 1145878

RESULT 4
 US-09-103-840A-1
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, Claire M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103.840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 US-09-103-840A-1

Query Match 40.1%; Score 632.6; DB 3; Length 4411529;

Best Local Similarity 68.4%; Pred. No. 9.4e-166;
 Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;
 140 AGAGGCGCACAGTGGCTGAAATCATGACGATATTCGCTCGCAAAATTCCTGACATCCCGCG 199
 1144551 AGGAGAAACCCAGTGGCGGATTCGAGCAGGTTAGGCGCCGAGAGATCCTCGATCCCGCG 1144610
 200 GTAAACCAACCGCTCGAGGCGAGAGTTTCTGTGATGACGGTCCCAAGGTTGCCAGGTG 259
 1144611 GCAACCGAGCGGTGGAGTGGCGCTTATCGACGGGACATTCGCCCGGCGCCGCG 1144670
 260 TTCATCCGCGGATCCACCGGCGTCCACGAGGCTCATGAGCTGCGTGCAGCGTGGCGCATC 319
 1144671 TGCGTGGGCGCTCGACCGGGGACACGAGCCGCTCGAGTTGGCGGACGCGCGCATC 1144730
 320 GTTACCTGGGCAAGGGGCTTTTGAAGGCAAGTTGAAACGTCACAAAGAAATTCGGCGACG 379
 1144731 GCTACGCGCGCAAAAGCGTGCAAAAGCCGTGAGGCTGTTCTTGATGAGATCGGCCCG 1144790
 380 AGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATCATCAAGCTTG 439
 1144791 CCGTTCATCGACTCAACGCGGACGACGAGCATTTGGTGCACGAGCGCTGGTGACCTAG 1144850
 440 ATGCGACCGCAACAAGTCCCGCTGGGTGCAAAACGCAATCCTTGGTGTTCATGGCTG 499
 1144851 ACGGACCCCGACAAAGTCCCGCTGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 1144910
 500 TTGAAAGGCTGCTGTGATTTCCGAGGCTCCACATGTTTCCGCTACATCGTGGAACAA 559
 1144911 TTGCAAGGCGGCGGCGATTCGCGGAGCTGCGTGTTCGTTATGTCGGGCGGCA 1144970
 560 ACCGACACGTTCTTCAGTTTCCAAATGATGAACATCAACAGTGGCGCTCAGCTGACT 619
 1144971 ACCGACACATTCGCGGTAACCGATGATGAACATCTCAACGCGGCGGACACGCCGATA 1145030
 620 CCGTGTGTGACGTTTCAGGAAATTCATGATCGCTCCAATCGGTGCGAGACCTTCTCTGAG 679
 1145031 CCGTGTGACATTCAGAGTTTCATGTTGGCGCAATGTTGGCGGCCAGCTTCTGTCAGG 1145090
 680 CTCTCGCAACGCGCGGAGGCTTACCAACGCTGAAGTCCGTCATCAAGAAAGAGGCC 739
 1145091 CGTTGGCTGGGCGCTGAGGTGTACCAACGCTCAAGTCGCTCTGAAAGAGAGGCGC 1145150
 740 TGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCGCTCCACCGTTCAGG 799
 1145151 TGTCCACCGGCTGGGCGAGGAGGCGCTTCGCGCGGATGTGCGCGGACCAACCGCG 1145210
 800 CTCTTGACCTTATCGTTGAGCAATCGAGAGGCTGGCTTTCACCCCGGCAAGGACATCG 859
 1145211 CGTTGGACCTGATCAGCGCGGCCATCGAGTGGCGGCTTCGACCGCGCGCGACGTG 1145270
 860 CTCTTGCTTGGAGTTGCTTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG 916
 1145271 CGTGGCCCTTGGACCGGCGGCCACCGAGTTCTTCAACGCGCACCGGCTAGCTCTCG 1145330
 917 AAGTGGCCAGCACTCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTGCTTGACCGCT 976
 1145331 AGGCAACACCCGTACCGCAGACAGATGACCGAGTTCTAGCGGGCTGCTCGCGGCT 1145390
 977 ACCCAATCGTCTCATCGAGACCCACTGCAAGGAAGATGACTGGAGGGTTTACACCAACC 1036
 1145391 ACCCGCTGGTGTGATCGAAGAGACCCACTGTCCGAAGACGATTTGGGACGCTGGGCGCGC 1145450
 1037 TCACCGCAACCATCGGCGACAAAGTTTCAGATCGTTGGCGGAGCTTCTTGTGACCAACC 1096
 1145451 TGACGGCTCGATCGGTGACCGGGTGCAATTCGTGGCGGACGACATCTTTGTCAACCAATC 1145510
 1097 CTGAGCGCTGAAGGAGGCGATCGCTAAGAAAGGCTGCCAATCCATCTCTGTTAAAGTGA 1156
 1145511 CCGAGCGGCTCGAGGAGGCGATCGAACGGGCGCTGGCAAAATGCGTTGCTGTAAGGTGA 1145570
 1157 ACCAGATCGGTACCTTCACCGAGACCTTTCAGCGTGTGACATGGCTCACCGCGCAGGCT 1216

Db 1145571 ACCAGATCGGAGCGTTGACCGAGACATCGACGCGGTCACTGCTGCTCACCACCGCGGAT 1145630
 Qy 1217 ACACCTCATGATGTCCACCGTTCCGGTGAGAGCCGAGGACACCACTTCTGACCTCG 1276
 Db 1145631 ACCGACGATGATCAGTCAACCGCAGTGCAGAGACGAGACACCATGATCGCCGACCTCG 1145690
 Qy 1277 CAGTGTCACTCAACTGTGGCCAGATCAAGACTGTGCTTCAGCAGGTTCCGACCGTGTG 1336
 Db 1145691 CGGTGGCCATCGGACGCGGCGAGATCAAGACGCGCGCCTGCTCGCAGTGAAGCGCTCG 1145750
 Qy 1337 CAAGAGTCAACACGCTTCTCCGCATCGAGCAGTGTGTTGGCGAGCGCGGCTCTACGCGAG 1396
 Db 1145751 CAAGATACACGCTGCTGCGGATCGAAGAGCGCTTGGCGAGCGCGCCCGCTACGCGG 1145810
 Qy 1397 GTCGACGCGCATCCCACGCTTTCAGGCTTAATAAAGCGCT 1439
 Db 1145811 GCGACCTGGCATTTCTCGGTTCGGTGCAGACGAATAGGT 1145853

RESULT 5

US-09-902-540-4780
 ; Sequence 4780, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 4780
 ; LENGTH: 1299
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-4780

Query Match 35.0%; Score 552; DB 4; Length 1299;
 Best Local Similarity 65.7%; Pred. No. 6.3e-145;
 Matches 842; Conservative 0; Mismatches 425; Indels 15; Gaps 2;
 Qy 152 TGGCTGAAATCATGCAGTATTCTGCTCGCGAAATTCGACTCCCGGCGGTAAACCAACCG 211
 Db 2 TGACCGAGATTTCAGATTCTGGCGCGCGAAGTGTCTGACTCCCGTGGCAACCCGACCG 61
 Qy 212 TCGAGGCGAGAGTTTCTCGATGACGGTTCCTCCAGCGTGTGCGAGGTGTTCCATCCGCGC 271
 Db 62 TGGAGGCGGAGTCCAGCTTCCGGGGGCGCTCGTGGCGGCGCGCGGTGCGGTCCGCTG 121
 Qy 272 CATCCACGCGGCTCAAGAGGCTCATGAGTGTGCTGAGCGTGGC---GATCGTACCTGG 328
 Db 122 CGTCCACCGGTGAGCATGAGGCCATCGAGTGTGCGCGACGCGGCAAGCATGCTACCTGG 181
 Qy 329 GCAAGGCGTTTGAAGGCACTTGAAGCGTCAAGAGTCAAGAGAAATCGGCGAGCTCGCTG 388
 Db 182 GCAAGGCGGTGCAAGAGGCGGTGAAGAACGTGGTCTGAGCTGCTCGCGCCCGCTGGTGG 241
 Qy 389 GCCTAGAGGCTGAGCATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGGCACCG 448
 Db 242 GCATGAGCGCGCGGATCAGTTCGGGTGGACCGAGCTGCTGAGCTGAGCGGACCG 301
 Qy 449 CCAACAAAGTCCCGCTGGGTGCAACGCAATCTTGTGTTTCCATGCTGTTCGAAAG 508
 Db 302 CCACCAAGGCAAGTGGGCGCAACGCAATCTCGCGGTGTCCATGCGCGCGCTCGCG 361
 Qy 509 CTGCTGTGATTCGCGAGGCTTCCCATGTTCCGCTACATCGTGGAGCAACCAACGACG 568
 Db 362 CCGCGGCGGACGCGACGCGGCTGCGGTGTACCGCTAGCTGAGGCGGCGTGCAGCGCGCA 421

Qy 569 TTCTTCAGTTCCTAATGATGAACATCATCAACCGTGGCGCTCACGCTGACTCCCGTGTG 628
 Db 422 CCCTGCGCGTCCCTCTGATGAACATCTCAACGCGCGCGCACGCGACACCGCGTGG 481
 Qy 629 ACCTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCA 688
 Db 482 ACCTGACAGGATTCATGATGCTGCGCGCGGTCTCTCTCTCGGAGGCGCTCGCGCT 541
 Qy 689 ACAGCGCGGAGGTCTACACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCG 748
 Db 542 GGGGCGCGAGGTGTTCCACGCGCTGAAGAAGATTCTCAAGGGCGCGAAGCTGGCACGG 601
 Qy 749 GACTTGGCGATGAGGCGGCTTCTCTCTCTCGGTGGCTCCACCGGTGAGGCTCTTGACC 808
 Db 602 GCGTGGCGGACGAGGCGGCTTACGCGCGGACCTGCGCGGCGAAAGAGAGCGCTGAAGC 661
 Qy 809 TTATCGTTGAGGCAATCGAAGGCTGCTTCAACCCAGGCAAGAGATCGCTCTTGTGTC 868
 Db 662 TCATCATGAGGCGCATCGACGCGGCTTCAAGGCGGTGAGAGCTCTTCTCTGGGCC 721
 Qy 869 TGGAGCTTGTCTCTCTGAGTTCTTCAAGGACGCGACCTACCACTT-----CG 916
 Db 722 TGGAGCTGCGCGGCGAGGAGTTCTTCCACAAGGCGAGCAAGATGACAGCTGAAGGCG 781
 Qy 917 AAGTGGCGGACACTCCGCGAGTGAAGTGGCAACGTTTACGCTGAGCTCTTGAACGCT 976
 Db 782 AGGCAAGGAGTACGACTCGACCGGCTGCTCGAGTACTACCGGGGCTCTCCGAGCGCT 841
 Qy 977 ACCCAATCGTCTCATCGAGGACCACTGCGAGGAGATGACTGGGAGGCTTACACCAACC 1036
 Db 842 ACCCATCATCTCCATCGAAGCGCATGCGGAGGATGACTGGGAGGCTTGAAGAACG 901
 Qy 1037 TCACCGCAACCATCGGCGCAAGGTTTCCAGATCGTTGGCGAGCTTCTTCTGCTACCAACC 1096
 Db 902 TCACCGAGCGCTGGGTTCCGCGCATGAGTTGGTGGGTGAGACCTCTTCTGCTACCAACG 961
 Qy 1097 CTGAGCGCTGAAGGAGGCGCATCGCTAAGAGGCTGCCAACTCCATCTGTTTAAAGTGA 1156
 Db 962 TGGAGCGCTGGGCGCGGATCGAGACGCGGCACTCCATCTCTGTTGAAGTGA 1021
 Qy 1157 ACCAGATCGGTACCTCACGAGACCTTCCAGCGTGTGAGATGGCTCACCGCGAGCT 1216
 Db 1022 ACCAGATGCGACCTTACGAGAGAGCTTTCAGCGCGTTCGCGATGCGCGACCGCGCGGT 1081
 Qy 1217 ACACCTCATGATGCTCCACCGTTCCGTTGAGACCGAGGACCAACCATCTGCTACCTCG 1276
 Db 1082 ACAGTCCGTCATGAGCACCGCTCCGCGGAGACGAGGACCAACCATCGCGACCTGG 1141
 Qy 1277 CAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCACTTCCGACGCTGTCG 1336
 Db 1142 CCGTGGCGTGGACTGCGGCGAGATCAAGACGCGTTCGCGTCCCGCTCCGACCGCGTGG 1201
 Qy 1337 CAAGATCAACAGCTTCTCGCATCGAGAGCTGCTTGGGAGCGCGGCTCTACGAG 1396
 Db 1202 CCAAGTCAACAGTTCGCGCATCGAAGCGAGCTGGGCGCGCGCGCTACGCTG 1261
 Qy 1397 GTCGACGCGCATTCGCCAGCTT 1418
 Db 1262 GCAAGTCCGTTTCCGCGCTT 1283

RESULT 6

US-09-902-540-1235/c
 ; Sequence 1235, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 1235
 ; LENGTH: 27903
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-1235

Query Match 35.0%; Score 552; DB 4; Length 27903;
 Best Local Similarity 65.7%; Pred. No. 3.1e-144;
 Matches 842; Conservative 0; Mismatches 425; Indels 15; Gaps 2;
 152 TGGCTGAATCATCGAGTATTCGCTCGCGAAATTCGACTCCCGGGTTAACCCACCG 211
 Db 1948 TGACCGAGATTCAGATTCGCGCGCGAAGTCTCGACTCCCGTGGCAACCGACCG 19889
 Qy 212 TCGAGGCGAGGTTTCTCGATGACGGTTCACCGGTGTCGACGGTGTTCATCCCGCG 271
 Db 19888 TGGAGCGGAGGTCCAGTTGCCGGGGCGCTGTCGCGCGCGCGCGTCCCGTG 19829
 Qy 272 CATCCACGGCGTCCACGAGGCTCATGAGTGGGTGACGTTGC---GATGCTACCTGG 328
 Db 19828 CGTCCACGGTGAAGTGAAGGCGCATCGAGTGGCGGACGCGCAAGCATCGCTACCTGG 19769
 Qy 329 GCAAGGGCGTTTGAAGCGAGTTGAAACGTCACGAGAAATCGGCGACGAGTCCGCTG 388
 Db 19768 GCAAGGGCGTGCAGAGGCGGTGAAGAACGTTGTCGACGCTGTCGCGCGCGCGTGGTG 19709
 Qy 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATCATCAAGCTTGATGGCACCG 448
 Db 19708 GCATGGAGCGCGGATCATGTTCCGGGTGGACACGCGATGCTGAGCTGACGCGCACCG 19649
 Qy 449 CCAACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTGCAAAAG 508
 Db 19648 CCACAAAGGCAAGTGGCGGCAACGCAATCTCGGGTGTTCATGGCGCGCTCGCG 19589
 Qy 509 CTGCTGCTGATTCGCGAGGCTCCCATGTTCCGCTACATCGGTGACCAACGACACG 568
 Db 19588 CCGCGGCGACGCGACGCGGTGCGCTGTACCGCTACGTGGGGGGGTGAGCGCGCA 19529
 Qy 569 TTCTTCCAGTTCCAAATGATGAATCATCAACGGTGGCGTCAAGCTGCTGCTGCTGGT 628
 Db 19528 CCCTGCGGTGCGCTGATGAATCATCTCAAGCGCGCGCACGCGACACCGCGTGG 19469
 Qy 629 ACCTTACGAAATTCATATGCTCCATCGGTGAGAGACCTTCTCTGAGGCTTCGCGCA 688
 Db 19468 ACCTGCGAGGATTCATGTTGTTGCGCGCGCTGCTCTCTTCCGCGAGGCGCTGCGCT 19409
 Qy 689 ACGCGCGGAGGTTTACACGACTGAAGTCCGTCATCAAGGAAAGGGCTGTCCACCG 748
 Db 19408 GGGCGCGGAGGTTTCAACGCTGGAAGATTCATCAAGGGCGGCAAGTGGCCACCG 19349
 Qy 749 GACTTGGCGATGAGGGCGGCTGCTCTTCCGCTCCAGCTCCACCGTGAAGCTTTGACC 808
 Db 19348 GCGTGGCGGACGAGGGCGGCTACGCCCGGACCTCGCGGCAAGGAGCGCTGAAGC 19289
 Qy 809 TTATCGTTGAGGCAATCGAAGAGCTGCTTCAACCGGCGGCGGCGGCGGCGCTGCTC 868
 Db 19288 TCATCATGGAGGCGCATCGACGAGGCGGCTTCAAGGCGGCTGAGCAGCTTCTCGGCC 19229
 Qy 869 TGGAGGTTGCTTCTCTGAGTTCTTCAAGGAGGCGCACTACCACTT-----CG 916
 Db 19228 TGGAGGTTGGCGGCGAGGAGTTCTTCAAGGAGGCGGCAAGTGAAGGCG 19169
 Qy 917 AAGTGGCGGAGCACTCCGAGCTGAGATGGAACAGTTTACGTGAGCTGTTGACCGCT 976
 Db 19168 AGGCAAGGAGTACGACTCGACCGGCTGCTGAGTACTACCGGGGCTCTCCGAGGCT 19109
 Qy 977 ACCCAATGCTTCCATCGAGGACCCACTGCGAGGAGATGACTGGAGGGTTACACCAAC 1036
 Db 19108 ACCCATCATCTCCATCGAAGAGCGGATGCGGAGGATGACTGGAGGGCTGGAAGAGC 19049

Qy 1037 TCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGGACGACTTCTTCGTACCAACC 1096
 Db 19048 TCACCGACGCGCTGGGTTTCGCGCATGCAAGTTGGTGGGTGACGACTTCTTCGTACCAACG 18989
 Qy 1097 CTGAGCGCCTGAAGGAGGCGATCGCTAAGAGGCTGCAACTCCATCCTCGTTAAGGTGA 1156
 Db 18988 TGGAGCGCCTGGGCGCGGCGCATCGAGACGGGCGGCAACTCCATCCTCGTTAAGGTGA 18929
 Qy 1157 ACCAGATCGGTACCTCACCAGACCTTCAGACGCTGTCGACATGGCTCAGCGCGCAGGCT 1216
 Db 18928 ACCAGATTGCGACCTGACGAGAGGTTTCAGCGCGTGGCGCATGCGCGCCGCTT 18869
 Qy 1217 ACACCTCATGATGTCACCGCTTCGCTGAGACCGGAGGACCAACCATTCGTGACCTCG 1276
 Db 18868 ACACGTCGCTCATGAGCACCGCTCCGCGGAGACGGAGGACCAACCATTCGCGACCTGG 18809
 Qy 1277 CAGTTGCACTCAACTGTCGCGCAGATCAAGACTGCTGCTCCAGCACGTTCCGACCGTGTG 1336
 Db 18808 CCGTGGCGCTGGACTGGCGGCGATCAAGACGGGTTCCGCTCCGCTCCGACCGCGTG 18749
 Qy 1337 CAAAGTACAAACGAGTTCCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTTACCGAG 1396
 Db 18748 CCAAGTACAAACGAGTTCCTGCGCATCGAAGCGAGCTGGGGCGCGCGCTACGCTG 18689
 Qy 1397 GTGCGAGCGCATTCGCCAGGCTT 1418
 Db 18688 GCAAGTCGGTTCGCCGCGCTT 18667

RESULT 7
 US-09-252-991A-4449
 ; Sequence 4449, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 4449
 ; LENGTH: 1335
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-4449

Query Match 33.8%; Score 533.6; DB 4; Length 1335;
 Best Local Similarity 64.8%; Pred. No. 9.5e-140;
 Matches 848; Conservative 0; Mismatches 439; Indels 21; Gaps 3;
 131 AGTTGCGATAGGAGGCGACAGTGGCTCAATCATGCAAGTATTCGCTCGGAAATTCG 190
 Db 26 ACTTCGAGGTGTTAAAGAAATGGCAAGATCGTCGACATCAAGGGCGCGTGGCTCTCG 85
 Qy 191 ACTCCCGCGGTAAACCAACCGTCGAGCGAGAGGTTTTCTTGGATGACGTTCCACCGGTG 250
 Db 86 ACTCCCGCGGCAACCTTACCGTTGAAGCGGACGCTGATCTCTGGACACGGCATCGTCGCA 145
 Qy 251 TCGCAGGTGTTTCATCGGCGCATTCACCGCGCTCCACGAGGCTCATGAGCTGGTGAAC 310
 Db 146 GCGCTCGCGCGCTTCGCTGCTTCACCGGTTCCCGCGAGGGCCCTCGAGCTGCGCATG 205
 Qy 311 GTGGCGA---TCGCTACCTGGGCAAGGGCGTTTTTGAAGGAGTGAAGAGCTCAACCAAG 367
 Db 206 GCGACAGAGCGCTTACCTGGGCAAGGGCGTGTGAAGCGGTGGCCCAATCAACGCGC 265
 Qy 368 AAATCGGCGAGGAGCTCGCTGGCTTAGAGGCTAGCGATGACGATGACGCTCATCGAAGCA 427

Db	266	CGATCCGGCAGCCTGTGCTGGCAAGGACGGCGCGACCAAGAAAGCCCTCGAACCCGGA	325
Qy	428	TGATCAAGCTTTGATGGCAACCCCAACAAGTCCCGCTGGGTGCAAAAGCAATCTCTGGT	487
Db	326	TGATCGAGCTGGACGGCACCGAGAACAAAGCCCAAGCTGGGCGCCAAACGGATCTCTCGGG	385
Qy	488	TTTCCATGGCTGTGTCGCAAGGCTGCTGTGATTCGCGAGGCGCTCCACTGTTCGCTAC	547
Db	386	TGTCCTCTGGCTCGCGCAAGCCCGCACAGGCCAAGGGCGTACCGCTGTACCGCGACA	445
Qy	548	TCGGTGGACCAACGCGAC-----ACGTTCTTCAGTTTCCAAATGATGAACATCA	595
Db	446	TCGCCGACCTCAACGGCACTCCGGCCAGTACTCAATGCCGGTCCGATGATGAACATCA	505
Qy	596	TCAACGCTGGCGCTCACGCTGACTCCGGTGTGACGTTTCAGGAATTCATGATCGCTCCAA	655
Db	506	TCAACGGCGGCGAGCATGCGCAACAACGTCGATATCCAGGAATTCATGCTCCAGCGG	565
Qy	656	TCGGTGCAGAGACCTTCTCTAGAGCTCTCCGCAACGGCGCGGAGGTCTACACGCACTGA	715
Db	566	TCGGCGCGAAGAACTTTCGGCAGGCGCTTGCGCATTGGGCGCGAGATCTTCCATCACCTCA	625
Qy	716	AGTCCGTCATCAAGGAAAGGGCTGTCCACCGACTTGGCGGATGGGGCGGCTTCGCTC	775
Db	626	AGGCGGTGCTGAAAGCCCGTGCTGAACACCGCGTCGGTGAAGAGGCGGCTTCGGCG	685
Qy	776	CTTCCGTCGGCTCCACCCGTCGAGGCTCTGTGACTTATCTGTTGAGGCAATCGAAGGCTG	835
Db	686	CGAACCTGTGCTCCAAACGAAGACCCCTGGCGGCCATCGCCGAGGCGCTCGAGAAGCGC	745
Qy	836	GCATTACCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTGCTTCTCTGAGTTCTTCA	895
Db	746	GCTACAAGCTGGGCGGACGACGTGACCTGGGCGCTTGGACTGCGGCTCCAGCGAGTCTTCA	805
Qy	896	AGGACGCACTTACCATTCGAAAGTGCGCAGCA-----CTCCGCACTGAGATGCAAA	949
Db	806	AGGACGCAAGTACGACTGGAAGCGAAGCAAGGTATTCGACGCGCCGCTTCGCGC	865
Qy	950	ACGTTTACGCTGAGCTCGTTGACGCGTACCCAACTCGTCTCCATCGAGGACCCATCGCAGG	1009
Db	866	ACTACCTGGCGCGCTGACCCAGCGCTACCCGATCATCTCCATCGAGGACGGCATGGACG	925
Qy	1010	AAGATGACTGGGAGGGTTACCAACCTCACCGCAACCATCGGCGCAAGAGTTCCAGATCG	1069
Db	926	AGTCCGACTGGGCGCGCTGGAAAGGCGCTGACGCAAGATCGGGCGCAAGGTCCAACTGG	985
Qy	1070	TTGGCGACGACTTCTTGTCACCAACCTGAGCGCTGAAAGGAGGATCGCTAAGAAGG	1129
Db	986	TCGGCGAGCACTGTTGTAACCAACCAAGATCTCTCAAGGAAGGACATCGAAGAGGCA	1045
Qy	1130	CTGCCAACTCCATCCTGGTTAAGGTGAACCAAGATCGGTAACCTCACCGAGACTTCGACG	1189
Db	1046	TCGGCAATTCGATCTCTGATCAAGTTCAACACAGATCGGTTTCCTCAACGAGACCTTGAGG	1105
Qy	1190	CTGTGCAATGGCTCACCGCGAGGCTACACTCGATGATGTCCGACCGTTCCGGTGAGA	1249
Db	1106	CCATCCAGATGGCCAAAGCGCCCGGCTATACCGGGGTGATCTCGCACCGCTCCGGCGGAAA	1165
Qy	1250	CCGAGGACACCACTATTGCTGACTCGGAGTTGCACTCAACTGTGGCCAGATCAAGACTG	1309
Db	1166	CCGAGGACTCGAACCATGCGCACTTGGCCGTGGGTAACGCGCGCGGTTCAGATCAAGACCG	1225
Qy	1310	GTGCTCCAGCAGTTTCGACCGGTGCGAAAGTACAACCACTTCCTCCGCAATCGAGCAGC	1369
Db	1226	GTTCCGTGTGCGCTCCGACCGCGTGTCCAAGTACAACCACTGCTGCGCATCGAAGAC	1285
Qy	1370	TGCTTTGGGACCGCGCGCTCTACGAGGTTCGACGGCATTCCTCCAGCT	1417
Db	1286	AACCTGGGCGCAAGCGCGTACCGTGTGTCGCGGAAATTCGCGGGCT	1333

US-09-252-991A-4158/c
; Sequence 4158, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4158
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4158

Query Match 33.8%; Score 533; DB 4; Length 1347;
Best Local Similarity 65.1%; Pred. No. 1.4e-139;

Qy	146	CCACAGTGGCTGAAATCATGCACTGATATTCGCTCGCGAAATATCTTCGACTCTCCGCGGGTAACC	205
Db	1343	CAAGAAATGGCAAGAATCGTCACATCAAGGGCGCTGAGGTCTCTGGACTCCCGCGGCAACC	1284
Qy	206	CAACCGTTCGAGGCAGAGGTTTTCTCGATGACGGTTCCACCGTGTTCGACAGGTGTTCAT	265
Db	1283	CTACCGGTTGAAGCGGACGTAATCTCGACAAAGGCATCGTTCGCGACGCGCTCGCGCGCTT	1224
Qy	266	CCGCGCATCCACCGGCGTCCACGAGGCTCATCAGCTGCGTGAACGCTGGCGA--TCGCT	322
Db	1223	CCGGTGCTTCCACCGGTTCCCGGAGGCGCTCGAGCTGCGGATGGCGAACAGAGCGGTT	1164
Qy	323	ACCTGGGCAAGGGCGTTTGAAGGCAAGTTGAAACGCTCAACGAAGAAATCGGCGACGAGC	382
Db	1163	ACCTGGGCAAGGGCGTCTGAAAGCCGTGSCCAACATCAACGGGCCGATCCGCGACTGC	1104
Qy	383	TCGGTGGCTAGAGGCTGACGATCAGGGCCCTCATCGACGAGCAATGATCAAGCTTTGATG	442
Db	1103	TGCTGGGCAAGGACGGCGCGACAGAAAGCCCTCGACACCGCGATCGAGCTGGACG	1044
Qy	443	GCACGCCAACAAAGTCCCGCTCGGGTGCAAAAGCAATCTCTTGTTGTTTCATCGCTGTTG	502
Db	1043	GCACCGAGAACAGGCCAAGCTGGGCGCAACCGGATCTCTGCGGGTGTCCCTGGCTGCCG	984
Qy	503	CAAAGGCTGCTGTGATTCGCGAGGCGCTCCCACTGTTTCGCTACATCGGTGGACCAACG	562
Db	983	CCAAGGCCGCGCACAGGCCAACGGCGTACCGCTGTACGCGCACATCGCGACCTCAACG	924
Qy	563	CAC-----ACGTTCTTCAGTTCCAAATGATGAACATCATCAACGGTGGCGCTC	610
Db	923	GCACCTCCGCGCCAGTACTCCATCGCGGTCCGATGATGAACATCATCAACGGCGGCGAC	864
Qy	611	ACGCTGACTCCGGTGTGAGTTTCAGGAATTCATGATCGTCCAAATCGGTGCAGACCT	670
Db	863	ATCGGACAAACGTCGATATTCAGGAATTCATGTTTCAGCCGGTTCGGCGCGAAGAACT	804
Qy	671	TCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTTCGTCATCAAGG	730
Db	803	TCGCGGAGGCGCTGCGCATGGGGCGCCGAGATCTTCATCACCCTCAAGGCCGTCTGNAGG	744
Qy	731	AAAAGGCGCTGTCAACCGGACTTTGGCGATGAGGGCGGCTTCGCTCCTTCGTCGGGTCCA	790
Db	743	CCCGTGGCGCTGAACACCGCCGTCCGTGACGAAGGGCGCTTCGCGCGCAACCTGTCGTCCA	684
Qy	791	CCCGTGAAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGAGGCTGGCTTCAACCCAGGCA	850
Db	683	ACGAAGA CGCCTCGGCGCCATCGCCGAGGCGCGTCGAGAGGCGCGGTACAAGCTGGCGG	624

QY 851 AGGACATCGCTCTGCTCTGACGCTTCTCTCTGAGTCTTCAAGACGGCACCTACC 910
Db |||||
QY 623 ACGACGTGACCTCGCTCGCTGACCTGCGCTCTCAGCGAGTCTTCAAGACGGCAAGTACG 564
Db |||||
QY 911 ACTTCGAAGGTGGCCAGCA-----CTCCGAGCTGAGATGGCAAAAGTTTACGCTGAGC 964
Db |||||
QY 563 ACCTGGAAGGGAAGGCAAGGTAATCGACGCGCGGTTTCGCCGACTACCTGCGCCGCC 504
Db |||||
QY 965 TCGTTGAGCGGTACCCCAATCGCTCTCCATCGAGAGACCCATCGCAGAAAGATGACTGGGAGG 1024
Db |||||
QY 503 TGACCCAGCGCTACCCCGATCATCTCCATCGAGGACGGCATGGACGAGTCCGACTGGGCCG 444
Db |||||
QY 1025 GTTACACCAACTCACCGCAACCATCGGCAACAGTTTCAGATCGTTTGGGACGACTTCT 1084
Db |||||
QY 443 GCTGGAAGGCTGACCGACAAGATCGGCGCAAGGTCCAACTGGTCGGCGACACCTGT 384
Db |||||
QY 1085 TCGTCACCAACCCCTGAGCGCTGGAAGGAGGCAATCGCTTAAGAAGGCTGCCAACTCCATCC 1144
Db |||||
QY 383 TCGTGACCAACACCAAGATCTCAAGGAGGATCGAGAGGGCATCGGCAATTCGATCC 324
Db |||||
QY 1145 TGGTTAAGGTGAACAGATCGGTACCCCTACCGAGAGCTTCGAGCGTGTGACATGCTC 1204
Db |||||
QY 323 TGATCAAGTTTCAACAGATCGGTTTCGCTCACCGAGACCTCGAGGCCATCCAGATGCCA 264
Db |||||
QY 1205 ACCGCGAGGCTACACCTCCATGATGTCACACCGTTCCGCTGAGACCGAGGACACACCA 1264
Db |||||
QY 263 AGCGCGCGGCTATACCGCGGTGATCTCGCACCGCTCCGCGGAACCGAGGACTCGACCA 204
Db |||||
QY 1265 TTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTTGGTGTCCAGACGTT 1324
Db |||||
QY 203 TCGCGGACCTGGCGTGGTACCGCCCGGTGATCAAGACCGGTTTCGCTGCGGCT 144
Db |||||
QY 1325 CCGACGCTGTCGAAAGTACAAACAGCTTTCGCGCATCGAGCAGCTGTTGGGACGCG 1384
Db |||||
QY 143 CCGACCGCGTGTCAAAGTACAAACAGTTGCTGCGCATCGAAGAGCAACTGGGCGCCAAG 84
Db |||||
QY 1385 GCGTCTACGAGGTTCGAGCGCATTCACAGCT 1417
Db |||||
QY 83 CGCGTACCGTGTGCGCGGGAATTCGCGCT 51
Db |||||

RESULT 9
US-09-489-039A-2890
; Sequence 2890, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2890
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2890

Query Match 30.5%; Score 481.2; DB 4; Length 1380;
Best Local Similarity 62.5%; Pred. No. 5.3e-125;
Matches 813; Conservative 0; Mismatches 463; Indels 24; Gaps 3;

QY 140 AGGAGCCACAGTGGCTGAATCATGACAGTATTCGCTCGCGAAATTCGACTCCGCG 199
Db |||||
QY 71 AGGAAAACCTTAATGTCCAAAATCGTTAAAGTCATCGGTGCGTGAATCATCGACTCCGCTG 130
Db |||||
QY 200 GTAACCAACCGTCGAGCGCAGAGTTTCTCGATGACGTTCCCAAGGTGTCGAGGTG 259
Db |||||
QY 131 GTAACCCGACTGTGTAAGGCCGGAAGTACACCTGGGAAGTGGTTTCGCTGGTATGSCAGCTG 190
Db |||||

QY 260 TTCCATCCGCGGCATCCACCGGCTTCAACGAGGCTCATGAGCTGCGTGAAGCGGTGGCA-- 317
Db |||||
QY 191 CTCGCTACGCTGCTTCTACTGGTTTCGCGCAAGCGCTGGAACTGCGCGATGGGACAAAT 250
Db |||||
QY 318 -TCGCTACCTTGGGCAAGGGGTTTGAAGGCGAGTTGAAAACGTTCAACGAAGAAATCGCGC 376
Db |||||
QY 251 CCGGTTTCTGGTTAAAGGCGTAACCAAGCTGTTCGTCGGTTAAACGGCCGATCGCTC 310
Db |||||
QY 377 ACAGGCTCGCTGCGCTAGAGGCTGACGATCAGGCGCTTCATCGACGAAGCAATGATCAAGC 436
Db |||||
QY 311 AGCAATCTCTGGGCAAAAGATGCTAAAGATCAGGCTGCGATCGCAACAGATCATGATGACC 370
Db |||||
QY 437 TTGATGGCACCGCAACAAAGTCCGCTCGGTGGAAGCAATCTTGGTTTTCATG 496
Db |||||
QY 371 TGGACGGTACTGAAAAAATACTAACTTCGTTGCGTGGAAACGCTATCTCTGGCCGTTTTC 430
Db |||||
QY 497 CTGTGCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAC 556
Db |||||
QY 431 CGAACGCCAAGCAGCAGCTGCGTCAAGGCTTGGCGCTGTAGCGGCATCTGCTGAAC 490
Db |||||
QY 557 CAAACGC-----ACAGTTCTTCCAGTTCCAATGATGAACATCATCAACGGTG 604
Db |||||
QY 491 TGAACGGCACTCCGGGCAAAATCTCCATGCGGTTCCGATGATGAACATCATCAACGGCG 550
Db |||||
QY 605 GCGCTACGCTGACTCGGTGTTGAGGTTTCAAGAAATTCATGATCGCTCCAATCGGTGCG 664
Db |||||
QY 551 GTGAGCACCGTGAACAACAGCTCGACATCCAGAAATTCATGATTCAGCCGCTTGGCGCG 610
Db |||||
QY 665 AGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACACGCACTGAAGTCCGTCA 724
Db |||||
QY 611 CGACCTTGAAGAAGCAGTAGCATGGGTTCTGAGTTTCCATCACTGCGCAAGTGC 670
Db |||||
QY 725 TCAAGGAAAAGGCGCTGTCCACGGACTTGGCGATGAGGCGGCTTGGCTCTTCCGTGCG 784
Db |||||
QY 671 TGAAGTCCAAAGGCATGAACACTGCGCTTGGTACGAAGCGGCTACGCGCCGAACCTGG 730
Db |||||
QY 785 GCTTCCACCGTGAAGGCTTTCACCTTATCGTTGAGGCAATCGAAGGCTGGCTTACCC 844
Db |||||
QY 731 GTTCCAACGCCGAAGCGCTGGCTGTATCGCTCAAGCGGTTAAAGCTGCGAGGCTACGAGC 790
Db |||||
QY 845 CAGGCAAGGACATCGCTTTCGCTCTGGAAGTTCCTCTGAGTTTCTTCAAGACGGCA 904
Db |||||
QY 791 TGGGCAAGACATACCTTGGCGATGAGCTGCGCGGCTCTGAATCTTCAAGACGGTA 850
Db |||||
QY 905 CTTACCACTTCGAAGGTG-----GCCAGCACTCCGCGCTGAGATGGCAAAACGTTT 955
Db |||||
QY 851 AATACGTTCTGGCTGGCGAAGCAACAAAGCGTTTCACTCTGAAGAGTTTCACTCACTTC 910
Db |||||
QY 956 ACCTGAGCTCGTTGAGCGGTACCCATCGTCTCCATCGAGGACCCCACTCGCAGGAAGATG 1015
Db |||||
QY 911 TGAAGAGCTGACCAACAGTACCGCATCCGATCGTCTCCATCGAAGATGGTCTCGAAGATCTG 970
Db |||||
QY 1016 ACTGGGAGGTTTACACCAACCTCACCGCAACCATCGCGCAACAAAGTTTCAGATCGTTGGCG 1075
Db |||||
QY 971 ACTGGGAAGGTTTCGCTTACCAGCTAAAGTACTTGGCGCAACAAATCCAGCTGGTTGGTG 1030
Db |||||
QY 1076 ACGACTTCTGCTACCAACCTTGAGCGCTGGAAGGAGGCGCATCGTGAAGAGGCTGCCA 1135
Db |||||
QY 1031 ACGACCTGTTCGTAACCAACACCAAGATCTCTGAAGAAGGCATCGAAAAAGGCATCGCTA 1090
Db |||||
QY 1136 ACTCCATCTCTGGTTAAGGTGAACAGATCGGTACCTTCCAGAGAGCTTCAGCGCTGCTCG 1195
Db |||||
QY 1091 ACTCCATCTCTGATCAAGTTCAACAGATCGGTTCTCTGACCGCAAACTCTGGCTGCTATCA 1150
Db |||||
QY 1196 ACATGGCTACCGCGCAGGCTACACCTTCCATGATGTCCCAACCGTTCCGTTGAGACCGAGG 1255
Db |||||
QY 1151 AGATGGCAAGACGCTGGCTACACCGCTGTTATCTCTACCGCTTCTGGCGAAATCGAAG 1210
Db |||||
QY 1256 ACACCAACCTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCTC 1315
Db |||||
QY 1211 ACGTACCAATTCGCGACCTGGCTGTTGGTACCGCTGACGCGCAGATCAAAAAGGTTCTA 1270
Db |||||
QY 1316 CAGCAGGTTCCGACCGGTTCGCAAAAGTACAACCAAGCTTCTCCGATCGAGCAGCTGCTTG 1375
Db |||||

Db 1271 TGAGCGCTTGCACCGCTTCTAAATACAAACCAAGCTGATTCGTATCGAAGAGCGCTGG 1330
 Qy 1376 GCGAGCGCGCGCTTACGACGCTCGCAGCGCATTCGCCAG 1415
 Db 1331 GTGAGCAAGCGCGTTCACCGTTCGTAAGAGATCAAGG 1370

RESULT 10

US-09-818-780-22
 ; Sequence 22, Application US/09818780
 ; Patent No. 6677146
 ; GENERAL INFORMATION:
 ; APPLICANT: McHenry, Charles
 ; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME
 ; FILE REFERENCE: 1794.0030004
 ; CURRENT APPLICATION NUMBER: US/09/818,780
 ; CURRENT FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: US 60/192,736
 ; PRIOR FILING DATE: 2000-03-28
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 2363
 ; TYPE: DNA
 ; ORGANISM: Thermus thermophilus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Includes 5' UTR and 3' UTR
 US-09-818-780-22

Query Match 28.5%; Score 449.2; DB 4; Length 2363;
 Best Local Similarity 64.3%; Pred. No. 7.1e-116;
 Matches 692; Conservative 0; Mismatches 378; Indels 6; Gaps 1;

Qy 160 ATCATGACGATTCGCTCGCGAAATTCCTCGACTCCCGCGGTAAACCCACCGTCGAGCA 219
 Db 1279 ATCTCGCGCTCGCGGACGCGAGGTTTGGATTCACAGGGCTTTCCACCGTAGAGCG 1338
 Qy 220 GAGGTTTCTCGGATGAGTTCCACGCTGTCGAGGTTCATCCGGGCGATCCACC 279
 Db 1339 GAGGTGAGCTGGAAGGGGGGCCAGGGGCCATGGTGCCCTCCGGGGCTCCACC 1398
 Qy 280 GCGCTCCACGAGGCTCATGAGCTGCTGACGCTGCGGATCCCTACCTGGGCAAGGGGCTT 339
 Db 1399 GGAACCCACGAGGCTTGGAGCTCAGGACGCGGCGGACGCTACCTGGGCAAGGGGCTG 1458
 Qy 340 TTGAAGCAGTTGAAACGTCACGAAATTCGCGACGAGCTCGCTGGGCTTAGAGCT 399
 Db 1459 CGCGGGCGGTGGAGAACGTCACGAGCGCATCGCCCGAGCTCGTCGGCATGGAGCC 1518
 Qy 400 GACGATCAGCGCTCATCGACGAGCATGATGATGATGATGATGATGATGATGATGATGAT 459
 Db 1519 CTGACACGAGGAGGGGTGAGCGGCGCATGCTGAGCTGGACCGGACCGGCAACGAGCC 1578
 Qy 460 CGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTGCAAGGCTGCTGCTGAT 519
 Db 1579 AACCTGGAGGAGGAGCGCTCTCGCGTCTCCCTGGCCGTGGCCCGGCGGCGCGAG 1638
 Qy 520 TCCGAGCGCTCCCACTGTTCCGCTACATCGGTGAGCAACGACGACAGTTCCTCCAGTT 579
 Db 1639 GCGCTGGGCTGCGGCTTTACCGCTACCTGGGCGGGTCCAGGGGCTCACCTGCGCGTG 1698
 Qy 580 CCAATGATGACATCATCAAGCGTGGCGCTCACGCTGACTCCGCTGTTGAGCTTCAGGAA 639
 Db 1699 CCCCTCATGACGTCATCAACGCGGGGGAAGCACCGGAGTTCAGGCTTCAGGAG 1758
 Qy 640 TTATGATCGCTCCCAATCGGTGAGAGACCTTCTGAGGCTCTCCGCAACGCGCGCGAG 699
 Db 1759 TTATGCTGGTGGCGGGGGGAGAGCTTCGCGGAGGCTTCAGGATCGGGGCGAG 1818
 Qy 700 GTCTACCAACGACCTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGAT 759

Db 1819 GTCTTCCACACCCCTCAAGGGCGTCTCAAGAGAGGGCTACAGCAACCAAGTGGGGAC 1878
 Qy 760 GAGGGCGGCTTCGCTTCCTTCGCTCGGTCCACCGTGAGGCTCTTGACCTTATCGTTGAG 819
 Db 1879 GAGGAGGCTTCGCGGCGGACCTCAGGAGCAACAGGAGGCGGTGGAGCTTTTGTCTCTC 1938
 Qy 820 GCAATCGAAGAGGCTTTCACCCAGGAGGACATCGCTCTTGTCTCTGAGGCTTGTCT 879
 Db 1939 GCAATTCAGCGGGCGGGGTACACCCCGGCGCAGGAGGTCTCCCTGGCGCTTGAACCGGCC 1998
 Qy 880 TCCTCTGAGTCTTCAAGGAGCGGACCTACACCTTCGAGGTG-----GCCAGCACTCC 933
 Db 1999 ACAGGAGGCTTTACCGGAGCGGAGTACACCTGGAAGGGAGGCAAGGTCTCTCTCC 2058
 Qy 934 GCAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTCGCGGTACCCCAATCGTCTCCATC 993
 Db 2059 TCGAGGAGATGCTGGGCTTCTGGGAGGCTGGGTGGAGAGTACCCCATCCGCTCCATT 2118
 Qy 994 GAGGACCCACTGCGAGGAGATGACTGGGAGGGTTACACCAACCTCACCGCAACCATCGGC 1053
 Db 2119 GAGGACGCTTTCGCGAGGAGGCTGGGAGGGTGGCGGCTTCTCACCGAGCGCTGGGG 2178
 Qy 1054 GACAGGTTGAGATCGTTGGGAGGAGCTTCTTCGTACCAACCTCGAGCGCTGAGGAG 1113
 Db 2179 GGAAGGTTCCAGCTCGTGGGAGGAGCTTCTTCGTACCAACCGCGGAAAGGCTCGGGCG 2238
 Qy 1114 GGATCGCTAAGAGGCTGCCAATCCATCCATCTGTTAAGGTGAACAGAGATCGGTACCCCTC 1173
 Db 2239 GGNATTGAGCGGGGGTGGCCACCGCATCTCTGTTCAAGGTGAACAGATCGGAGCCCTC 2298
 Qy 1174 ACCGAGACCTTCGACGCTGTGACATGGCTCACCGCGCAGGCTACACCTCCATCAT 1229
 Db 2299 TCGAAACCTTCGAGGCACTCCGCTGGCCAGGCTCGGGGTACAGGGCGGTGAT 2354

RESULT 11

US-08-781-802-9
 ; Sequence 9, Application US/08781802
 ; Patent No. 5969121
 ; GENERAL INFORMATION:
 ; APPLICANT: ALLEN, Larry
 ; APPLICANT: AIKENS, John
 ; APPLICANT: FONSTEIN, Michael
 ; APPLICANT: VONSTEIN, Veronika
 ; APPLICANT: DEMIRJIAN, David
 ; APPLICANT: CASADABAN, Malcolm
 ; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 ; STREET: 300 S. Wacker Drive 32nd Floor
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,802
 ; FILING DATE: 10-JAN-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/694,078
 ; FILING DATE: 07-AUG-1996
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/019,580
 ; FILING DATE: 12-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/009,704


```
;
; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA: US 60/001,995
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-E
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3545 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1397..2905
; OTHER INFORMATION: /note= "E019 sequence of longest
; open reading frame; upstream untranslated region not exact"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1397..2905
; US-08-781-802-9

Query Match 28.0%; Score 442.4; DB 2; Length 3545;
Best Local Similarity 62.8%; Pred. No. 7.2e-114; Indels 15; Gaps 3;
Matches 743; Conservative 0; Mismatches 426;

Qy 229 CTGGATGACGGTCCACGGTGTGCGCAGGTGTTCCATCCGGCGCATCCACGGCGTCCAC 288
Db 39 CGGGAAGAAGCGGTTTCGCCCGTGTAGTGCCAAAGCGCGTTTCGACGGGGAATAT 98

Qy 289 GAGGCTCATGAGCTGCGTACGGTGGC---GATCGCTACCTGGGCAAGGGGTTTGAAG 345
Db 99 GAAGCGGTGAATTTGCGTGACGGCGACAAAACCGCTACCTCGGCAAAAGGGGTGCTCAA 158

Qy 346 GCAGTTGAAACGTCACGAAGAAATCGGCGACGAGCTCGTGGCTAGAGGCTGACGAT 405
Db 159 GCGGTTGGAACGTCACGAAGTGTGCTCCGGAATCATCGGCTTAGAAGTACGTAT 218

Qy 406 CAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGGCACGCGCAACAAAGTCCCGCTG 465
Db 219 CAAGTGGCGATCGACCGCGCTTGATTGAACCTTGACGGCACGGAACAAAGAAAGCTT 278

Qy 466 GGTGCAACGCAATCCTTGGTGTTCATGGCTGTGCAAGGCTGCTGCTGATTCGCA 525
Db 279 GGGCGGAATGCTATTATTAGGCGTGTGCTGCGCTGCGCTGCGCTGCGGCTGATGAGCTT 338

Qy 526 GGCTCCACATGTTCCGCTACATCGGTGGACCAAGCACAGTCTTCCAGTTCCAATG 585
Db 339 GCGTTCGCTGTACCAATATCTGGCGGCTTTAACGCTTAACGCTGCCCTGTAACGATG 398

Qy 586 ATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGACGTTGACGAAATTCATG 645
Db 399 ATGAACATTTTAAACGGCGGCGCATGCGGACAAACAGTGTGACATTTCAAGAATTCATG 458

Qy 646 ATCGCTCCAATCGGTGCAGAGACCTTCTCTAGGCTCTCCGCAACGGCGGAGGTCTAC 705
Db 459 ATCATGCGGTGCGTGGGAAAGCTTCGTGAAGCGCTGCGCATGGGTGCAGAAATTTTC 518

Qy 706 CACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACGGACTTGGGGATGAGGCG 765
Db 519 CATAGCTTAAAGCTGTGTTAAAGCGAAAGGCTACACACGCTGTGCTGCGTGAAGAGGC 578

Qy 766 GGCTTCGCTCCTTCGTCGGGTCCACCGTGAGGCTCTTGACCTTATGTTGAGGCAATC 825
Db 579 GGATTGCTCGAACTTAAATCGAACGAAGAAGCGCTGCAACGATCATTTGAAGCGATC 638
```

```
Qy 826 GAGAAGGCTGGCTTCACCCCAGGAGACATCGCTTCTGCTCTGGACGTTGCTTCTCT 885
Db 639 GAAAAAGCGGCTACAAACAGGCAACAGCGCTATGCTCGCTATGGACGTTGCTTCGTCG 698
Qy 886 GAGTCTTTCAAG-----GACGGCACCTACCACTT-----CGAAGGTGGCCAGCACTCC 933
Db 699 GAGCTGTACAAACAAGAGATGGCAATATCATTTGGAAGCGCAAGGCGTGGTCAAAACA 758
Qy 934 GCAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACCGCTACCCAAATCGTTCATC 993
Db 759 TCAGAAGAAATGTTGCTGCTATGAAGAGCTTGTGTCGAAATATCCGATCATCTCGATC 818
Qy 994 GAGACCCACTGCAGGAAGATGATGGGAGGGTTTACACCAACCTCACCAGCAACCATCGGC 1053
Db 819 GAAGACGGACTTGACGAAATGATGGGAAAGGCCATAAAGTCTTACTGAGCGCTTTGGC 878
Qy 1054 GACAAGGTTTCAGATCGTTGGCGACGACTTCTTGTTCACCAACCTGAGCGCTGAAGGAG 1113
Db 879 CACAAGTGCAGCTCGTGGTGAGCACTTGTGTAACGAACAGAAACAACTGGCCGAA 938
Qy 1114 GGATCGCTAAGAAAGGCTGCCAACTCCATCCTCGTTAAGGTGAACCAAGATCGGTACCTC 1173
Db 939 GGCATTGAAAAGCGGTCGGCAACTCGATTTTAATTAAGTGAACCAAAATCGGTACACTG 998
Qy 1174 ACCGAGACCTTCGACGCTGCGACATGGCTCACGGCGAGGCTACACCTCCATGATGCTC 1233
Db 999 ACGGAAACGTTTCGATGCTTGAAGTGGCCAAACGGCGCGGTACACGGCGGTTGTGTCG 1058
Qy 1234 CACGTTTCGGTGAAGCCGAGGACACCAACCATTTGCTGACCTCGCAGTTGCACTCAACTGT 1293
Db 1059 CACGTTTCGGTGAACCGGAAGACAGCACGATTTGCCGATATCGTGTGCAACAAACGCT 1118
Qy 1294 GGCCAGATCAAGACTGGTGTCTCCAGCACGTTCCGACCGTGTCCAAAGTACAAACGACTT 1353
Db 1119 GGCCAAATCAAAACGGGAGCACCGTTCGCTACGACCGCTCGCAAAATACAAACGACTG 1178
Qy 1354 CTGCGCATCGACGAGCTGCTTGGGAGCGCGGCTCTACGAGG 1397
Db 1179 CTGCGCATTAAGACGAACACTTGGCCACACGGCTATTTTACCAAG 1222
```

RESULT 12

```
US-09-583-110-1156
; Sequence 1156, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1156
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-1156
```

```
Query Match 27.5%; Score 433.2; DB 4; Length 1305;
Best Local Similarity 59.9%; Pred. No. 1.7e-111;
Matches 776; Conservative 0; Mismatches 493; Indels 27; Gaps 2;

Qy 160 ATCATGACGATTTTCGCTCCGCAAAATTCGATCTCCGCGGTAAACCAACGTCGAGCA 219
Db 10 ATTACTGATGTTTACGCTCGGAAGTCTTAGACTCAGCGGTAAACCAACACTTGAAGTA 69
```



```
Db 2884 CGTTACGGTGGTCTTTGGTACACAAAAAGCTGTTGACAAACGTAACAAACATCATTTGCTGAA 2825
Qy 379 GAGTCGCTGGCCTAGAGGCTGACGATCAGGCGCTCATCGACGAAAGCAATGATCAAGCTT 438
Db 2824 GCTATCATTTGGCTACGATGATCGTGTCAACAAAGCTATTGACCGTGCTATGATCGCACTT 2765
Qy 439 GATGGCAGCGCAACAGTCCCGCTGGGTGCAACGCAATCCTTTGGTGTTCATGGCT 498
Db 2764 GACGGTACTCTTACAAGAAATTTGGGTGCGAATGCAATCCTCGGTGTGTCTATCGCT 2705
Qy 499 GTTCAAAGGCTGCTGTGATTTCCGAGGCGCTCCACTGTTTCGCTACATCGGTGGACCA 558
Db 2704 GTAGCTCGTGTGCTGCTGACTACTCTTGAATCCACTTTACAGCTATCTTTGGTGGATTC 2645
Qy 559 AACGACACGCTTCTCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGAC 618
Db 2644 AACACTAAAGTTCTTCCAACTCCCAATGATGAACATCATCAACGGTGGTTCCTCACTCGAC 2585
Qy 619 TCCGGTGTGAGTTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAG 678
Db 2584 GCTCCAATCGCTTCCAAAGTTCATGATCTTGCGAGTTGGTGCGCCAACATTTAAAGAA 2525
Qy 679 GCTCTCCGCAACGGCGGAGGTCTACACGCACTGAAGTCCGTTCATCAAGGAAAGGCG 738
Db 2524 GCCCTTGGTTACGGTGTGAAATCTTCCAAGCTCTTAAGAAATCCTTAAATCAGTGGT 2465
Qy 739 CTGTCCACGGAATTTGGGATGAGGGCGCTTCGCTCCTTCGCTCGGCTCACCCGCTGAG 798
Db 2464 TTGGAATCTGCGTAGTGACGAAGGTGGATTCTGCTCTCTCTTTCGAAGGAATCGAAGAT 2405
Qy 799 GCTCTGACCTTATCGTTGAGGCAATCGAAGGCTGCTTCAACCCAGGCAAGGACATC 858
Db 2404 GGTGTGAAACTATCTTCTGCGAATGAAGCTCTGGATATGTACCAAGGTAAAGACGTA 2345
Qy 859 GCTCTGCTGCGAGTTGCTTCTCTGAGTTCTTCAAGGAGCGCACCTACCACTTCGAA 918
Db 2344 TTTATCGGATTTGACTGTGCTTTCATCAGAAATCTTACGATAAGAACGTAAGTTTACGAC 2285
Qy 919 -----GGTGGCCAGCACTCCGCGAGCTGAGATGGCAAAAGCTT 954
Db 2284 TACACTAAATTTGAAGGTGAAGGTGCTGCTGTTGCTATCTGCAGAAACAAATCGACTAC 2225
Qy 955 TACGCTGAGCTCGTTGACGCTACCAATCGCTTCCATCGAGGACCCACTCGCAAGAGAT 1014
Db 2224 CTTGAAGAAATTTGGTTAAACAAATACCAATCATCACTATTGAAGATGGTATGGATGAAGAC 2165
Qy 1015 GACTGGGAGGTTTACACCAACTCACCGCAACCATCGGCGCAAGAGTTTCAGATCGTTGGC 1074
Db 2164 GACTGGGATGCTTGGAAAGCTTTACTGAACGCTCTTGGTAAGAAAGTACAACTTTGTTGT 2105
Qy 1075 GACGACTTCTTGGTACCAACCTGAGCGGCTGAGGAGGCGATCGCTTAAGAGGCTGCC 1134
Db 2104 GACGACTTCTTCTGTAACCAACACTGACTACCTTGCACGCTGTATCCAAAGAGGTGCTGCT 2045
Qy 1135 AACTCCATCTTGGTTAAGTGAACAGATCGGTACCTCCATCGAGACCTTCGACGCTGTC 1194
Db 2044 AACTCAATCTTATCAAAAGTTAAACAAATCGGTACTCTTACTGAACATTTTGAAGCTATC 1985
Qy 1195 GACATGGCTCACCGGAGGCTTACACCTTCCATGATGTCCACCGTTCGAGTGAACCGAG 1254
Db 1984 GAAATGGCTTAAGAGCTGGTTTACACTGCTGTTATCACACCGTTTCAAGTTGAAGCTGAA 1925
Qy 1255 GACACCACTTGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCT 1314
Db 1924 GATTCAACAAATCGCTGATTTGCAAGTTGCAAGTTAACGAGGACAAATCAAGACTGGTTCA 1865
Qy 1315 CCAGACGTTCCGACGCTGTCGAAGTACACAGCTTCTCCGATCGAGCAGCTGCTT 1374
Db 1864 CTTTCAGCTACAGACCGCATCGCTTAATACAAACCAATTTGCTTCGTATCGAAGACCAACTT 1805
Qy 1375 GCGACGCGGCGCTTACGCGAGGTGCGAGCGCATTTCCACGCTTTTCAGGCTAA 1428
Db 1804 GGTGAAGTAGCTGAATATCGTGGATTGAATCATTTCTACAACCTTTAAAAAATAA 1751
```

RESULT 14

```
US-09-107-532A-156
; Sequence 156, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1196
; SEQUENCE DESCRIPTION: SEQ ID NO: 156:
US-09-107-532A-156
```

```
Query Match 27.28; Score 429.6; DB 4; Length 1196;
Best Local Similarity 61.54; Pred. No. 1.6e-110;
Matches 731; Conservative 0; Mismatches 439; Indels 18; Gaps 2;

Qy 259 GTTCCATCCGCGCATCCACGGCGTCCACGAGGCTCATGAGGCTCATGAGCTGCGTACGGTGGC--- 315
Db 6 GTTCTTCAGTGCTTCAACTGGTGAATACGAGCAGTTGAATTCGTGACGGAGACAA 65
Qy 316 GATCGCTACCTGGCGAAGGGGCTTTTGAAGCGAGTTGAAACGTCACGAAAGAAATCGGC 375
Db 66 GCTCGTTATGCGGTAAAGGTGAACATAAGCTGTCCGACAAACGTAACATCATTTGCT 125
Qy 376 GACGAGCTCGTCCGCTAGAGGCTGAGATCAGCGCTCATCGACGAGCAATCATCAAG 435
Db 126 GAAGCAATCATTTGGCTTACGATGTTTCGTGACCAATGGCTATCGACAAGCTATGATCGCT 185
Qy 436 CTTGATGGCAGCGCAACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATG 495
```


Qy	1081	TTCTTCGTACCAACCTGAGCGCTGAGAGGGGATCGCTAAGAAGGCTGCCAATCC	1140
Db	393	CTGTTCTGTAACCAACACCAAGATCCTGAAAGAAGGCATCGAAAAAGGCATCGCTAACTCC	334
Qy	1141	ATCTGTGTTAAGGTGAACCCAGATCGGTACCTCACCGAGACCTTCGACGCTGTGACATG	1200
Db	333	ATCTGATCAAGTTCACACAGATCGGTTCCTGACCGAAGCTCTGGCTGCTATCAAGATG	274
Qy	1201	GCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGGTGAGACCGAGACAC	1260
Db	273	GCGAAAGACGCTGGCTACACCGCTGTTATCTCTCACCGTTCTGGCGAAACTGAAGACGCT	214
Qy	1261	ACGATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCTCCAGCA	1320
Db	213	ACGATTGCCGACCTGGCTGTGGTACCGCTGACGCCGAGATCAAAACTGGTTCATGAGC	154
Qy	1321	CGTTCCGACGCTGTCGCAAGTACAACAGCTTCTCGGCATCGAGCAGCTGTGGCGAC	1380
Db	153	CGTTCTGACCGGTTGCTAATAACAACAGCTGATTGTTATCGAAGAGCGCTGGGTGAG	94
Qy	1381	GCCGGGCTCTACGAGGTCGACGCGCATTTCCACG	1415
Db	93	CAAGCGCGTTCAACGCTCGTAAGAGATCAAAGG	59

Search completed: September 28, 2005, 23:33:36
 Job time : 294.5 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:02:19 ; Search time 6707 Seconds
(without alignments)
11400.377 Million cell updates/sec

Title: US2005728-947-1-1
Perfect score: 1578
Sequence: 1'ggctgggagataggtggtt.....ctcaagcaggagcgtgctt 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1:	gb_ba.*
2:	gb_hcg.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pat.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_sts.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6 AR490885	AR490885 Sequence
2	1578	100.0	1578	6 AX136862	AX136862 Sequence
3	1578	100.0	1578	6 BD014844	BD014844 Polynucle
4	1578	100.0	333150	1 AP005277	AP005277 Coryneb
5	1578	100.0	348475	1 BX927150	BX927150 Coryneb
6	1578	100.0	349980	6 AX127146	AX127146 Sequence
7	1576.4	99.9	1578	6 AR490886	AR490886 Sequence
8	1403.4	88.3	1405	6 AX763190	AX763190 Sequence
9	1396.4	88.5	1398	6 AX064945	AX064945 Sequence
10	1275	80.8	1275	6 BD163286	BD163286 Novel pol
11	1275	80.8	1275	6 AX121169	AX121169 Sequence
12	1036.4	65.7	300750	1 AP005217	AP005217 Coryneb
13	917.4	58.1	347625	1 BX248356	BX248356 Coryneb
14	734	46.5	276800	1 SC039315	SC039315 Streptomy
C 15	701.6	44.5	110000	1 AE016822	Continuation (18 o
C 16	698.4	44.3	296300	1 AP005035	AP005035 Streptomy
C 17	680.6	43.1	110000	1 AP006618	Continuation (51 o
C 18	657.6	41.7	303855	1 AE017230	AE017230 Mycobacte
C 19	642.4	40.7	30159	6 CQ363776	CQ363776 Sequence

20	642.4	40.7	110000	1 AE017283_05	Continuation (6 of
21	632.6	40.1	45191	6 AX704278	AX704278 Sequence
22	632.6	40.1	110000	1 AE000516_11	Continuation (12 o
23	632.6	40.1	327650	1 BX248337	BX248337 Mycobacte
24	632.6	40.1	349306	1 BX842575	BX842575 Mycobacte
25	588.4	37.3	329100	1 SME591787	AL591787 Sinorhizo
C 26	584.6	37.0	10029	1 AE014725	AE014725 Bifidobac
C 27	584.6	37.0	349980	6 AX492784	AX492784 Sequence
C 28	584.6	37.0	349980	6 AX492785	AX492785 Sequence
C 29	584.6	37.0	349980	6 AX553951	AX553951 Sequence
C 30	584.6	37.0	349980	6 AX553952	AX553952 Sequence
31	583.4	37.0	344050	1 MLEPRTN1	AE583917 Mycobacte
32	577.8	36.6	12058	1 AE009103	AE009103 Agrobacte
33	577.8	36.6	12407	1 AE008068	AE008068 Agrobacte
34	577	36.6	347660	1 AP002994	AP002994 Mesorhizo
C 35	576.4	36.5	10934	1 AE005846	AE005846 Caulobact
C 36	576.4	36.5	12709	1 AE014414	AE014414 Brucella
37	573.2	36.3	13776	1 AE009526	AE009526 Brucella
38	571.8	36.2	1290	1 AY181039	AY181039 Methyloba
C 39	571.4	36.2	349981	1 BX572602	BX572602 Rhodopseu
C 40	558	35.4	300600	1 AP005952	AP005952 Bradyrhiz
41	557.8	35.3	193050	1 AL646062	AL646062 Ralstonia
42	554	35.1	197615	5 BX470161	BX470161 Zebrafish
43	551.8	35.0	110000	1 AE017180_24	Continuation (25 o
44	551.8	35.0	110000	1 AE017180_25	Continuation (26 o
45	543.6	34.4	110000	1 AP006840_02	Continuation (3 of

ALIGNMENTS

RESULT 1

AR490885

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORIGIN

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 1578

/organism="unknown"

/mol_type="genomic DNA"

Query Match

Best Local Similarity

Matches 1578;

Conservative

Mismatches

Indels

Gaps

1578 bp DNA linear [PAT-15-MAY-2004]

Sequence 1 from patent US 6713289.

AR490885

AR490885

AR490885.1

GI:47258404

Unknown.

Unclassified.

1 (bases 1 to 1578)

Mockel, B., Pfeifferle, W., Hermann, T., Puhler, A., Kalinowski, J. and

Bathe, B.

Nucleotide sequences which code for the eno gene

Patent: US 6,713,289 A, 15 MAR 2004;

Location/Qualifiers

1. 1578

/organism="unknown"

/mol_type="genomic DNA"

Query Match

Best Local Similarity

Matches 1578;

Conservative

Mismatches

Indels

Gaps

1578 bp DNA linear [PAT-15-MAY-2004]

Sequence 1 from patent US 6713289.

AR490885

AR490885

AR490885.1

GI:47258404

Unknown.

Unclassified.

1 (bases 1 to 1578)

Mockel, B., Pfeifferle, W., Hermann, T., Puhler, A., Kalinowski, J. and

Bathe, B.

Nucleotide sequences which code for the eno gene

Patent: US 6,713,289 A, 15 MAR 2004;

Location/Qualifiers

1. 1578

/organism="unknown"

/mol_type="genomic DNA"

Query Match

Best Local Similarity

Matches 1578;

Conservative

Mismatches

Indels

Gaps


```
Db 241 TCCACGGTCTCGAGGTGTTCCATCCGGCGCATCCACCGCGTCCACGAGGCTCATGAG 300
Qy 301 CTGCGTGAACGGTGGCGATCGCTACCTGCGGCAAGGGCGTTTGAAGGCGATTTGAAAACGTC 360
Db 301 CTGCGTGAACGGTGGCGATCGCTACCTGCGGCAAGGGCGTTTGAAGGCGATTTGAAAACGTC 360
Qy 361 AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGGCTCATCGAC 420
Db 361 AACGAGAAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGGCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGTATGGCAGCGCCAAAGTCCCGCTCGGTCGAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGTATGGCAGCGCCAAAGTCCCGCTCGGTCGAAACGCAATC 480
Qy 481 CTTGGTGTTCATGGCTGTGTAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
Db 481 CTTGGTGTTCATGGCTGTGTAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
Qy 541 CGCTACATCGGTGGACCAACGACACGCTTCTCCAGTTCGAATGATGACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAACGACACGCTTCTCCAGTTCGAATGATGACATCATCAAC 600
Qy 601 GGTGGCGCTCACGCTGACTCGGTGTTGACGTTTACGAAATTCATGATCGCTCCAAATCGGT 660
Db 601 GGTGGCGCTCACGCTGACTCGGTGTTGACGTTTACGAAATTCATGATCGCTCCAAATCGGT 660
Qy 661 GCAGAGACTTCTCTGAGGCTTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTCC 720
Db 661 GCAGAGACTTCTCTGAGGCTTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTCC 720
Qy 721 GTCATCAAGAAAGGCGCTTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGAAAGGCGCTTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Qy 781 GTGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAGAGGCTTGGCTTC 840
Db 781 GTGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAGAGGCTTGGCTTC 840
Qy 841 ACCCAGGCAAGGACATCGCTCTTGGCTCTGAGGCTTCTCTGATGCTTCCAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTCTTGGCTCTGAGGCTTCTCTGATGCTTCCAGGAC 900
Qy 901 GGCACCTACCACTCGAAGTGGCCAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
Db 901 GGCACCTACCACTCGAAGTGGCCAGCACTCCGCACTGAGATGGCAACGTTTACGCT 960
Qy 961 GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
Qy 1021 GAGGGTTACCAACCTCAACGCAACCATCGGCAACGATGAGTTCAGATCGTTGGCGACGAC 1080
Db 1021 GAGGGTTACCAACCTCAACGCAACCATCGGCAACGATGAGTTCAGATCGTTGGCGACGAC 1080
Qy 1081 TTCTTCTGTCACCACTGAGCGCTGAGAGGGCATCGTAAAGGGCTGCCAATCC 1140
Db 1081 TTCTTCTGTCACCACTGAGCGCTGAGAGGGCATCGTAAAGGGCTGCCAATCC 1140
Qy 1141 ATCTCTGTTAAGGTGAACAGATCGGTACCTCAACGAGACCTTCGACGCTGTCGACATG 1200
Db 1141 ATCTCTGTTAAGGTGAACAGATCGGTACCTCAACGAGACCTTCGACGCTGTCGACATG 1200
Qy 1201 GCTCACCGCGAGGCTACACTCCATGATGTCACCGTTCGCGTGAGACGAGGACACC 1260
Db 1201 GCTCACCGCGAGGCTACACTCCATGATGTCACCGTTCGCGTGAGACGAGGACACC 1260
Qy 1261 ACCATTGCTGACCTCGAGTTGCACTCACTGTGGCCAGATCAGACTGTTGCTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGAGTTGCACTCACTGTGGCCAGATCAGACTGTTGCTCCAGCA 1320
Qy 1321 CGTTCCGACCGTTCGCAAAAGTACAACACAGCTTCTCCGCGATTCGAGCACTGCTTGGGAC 1380
```

```
Db 1321 CGTTCCGACCGTGTCCGAAAAGTACAACCAAGCTTCTCCGATCCAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCGGCGTCTACGCGAGTCCGACGCGCATTCGCCAGCTTTCAGGCTAAATAAAGCGCTT 1440
Db 1381 GCGGCGTCTACGCGAGTCCGACGCGCATTCGCCAGCTTTCAGGCTAAATAAAGCGCTT 1440
Qy 1441 TTTCGACGCCCGGTAACTCAAGGTTGCCGGGCGTGGTTGCCCTTACTCTGTTACTCGTGT 1500
Db 1441 TTTCGACGCCCGGTAACTCAAGGTTGCCGGGCGTGGTTGCCCTTACTCTGTTACTCGTGT 1500
Qy 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTCATAAAGGCTTGTTCGTCT 1560
Db 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTCATAAAGGCTTGTTCGTCTCT 1560
Qy 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 2
AX136862
LOCUS AX136862 1578 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1090998.
ACCESSION AX136862
VERSION AX136862.1 GI:14273213
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Moeckel,B., Pfeifferle,W., Hermann,T., Puehler,A., Kalinowski,J. and
Bathe,B.
TITLE Nucleotide sequences coding for the eno gene
JOURNAL Patent: EP 1090998-A 1 11-APR-2001;
Degussa AG (DE)
FEATURES
    Location/Qualifiers
        1..1578
            /organism="Corynebacterium glutamicum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1718"
            151..1428
                /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /db_xref="GI:14273214"
            /protein_id="CAC39860.1"

CDS
    100.0%; Score 1578; DB 6; Length 1578;
    Best Local Similarity 100.0%; Pred. No. 6.9e-288;
    Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
    1 GGCCTGGGGATATGGGTAGTTTTCGCCACATAATTTCAACTGATTCCTCATCGAAAACAAG 60
    1 GGCCTGGGGATATGGGTAGTTTTCGCCACATAATTTCAACTGATTCCTCATCGAAAACAAG 60
    61 TTCTGTCACAAATTTGGGTGTAGACGTTGATGAAGACATTTGATCAGTGAATTTCTAG 120
    61 TTCTGTCACAAATTTGGGTGTAGACGTTGATGAAGACATTTGATCAGTGAATTTCTAG 120
    121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
    121 TTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGTATTCGCTCGC 180
```

QY 181 GAAATTCCTGACTCCCGGGTAACCAACCGTTCAGGCGAGAGGTTTCTTGATGACGGT 240
 Db 181 GAAATTCCTGACTCCCGGGTAACCAACCGTTCAGGCGAGAGGTTTCTTGATGACGGT 240
 QY 241 TCCACAGGTTGCGAGGTTTCCATCCGGCGCATCCACCGCGTCACAGGCTCATGAG 300
 Db 241 TCCACAGGTTGCGAGGTTTCCATCCGGCGCATCCACCGCGTCACAGGCTCATGAG 300
 QY 301 CTGCGTGACGGTGGCGATCGCTACCTGGGCAAGGCGTTTGAAGGCGAGTTGAAAACGTC 360
 Db 301 CTGCGTGACGGTGGCGATCGCTACCTGGGCAAGGCGTTTGAAGGCGAGTTGAAAACGTC 360
 QY 361 AACGAAGAAATCGGCGAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
 Db 361 AACGAAGAAATCGGCGAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
 QY 421 GAAGCAATGATCAAGCTTGTATGGCAACCGCAACAAAGTCCCGCTGGGTGCAAGCGCAATC 480
 Db 421 GAAGCAATGATCAAGCTTGTATGGCAACCGCAACAAAGTCCCGCTGGGTGCAAGCGCAATC 480
 QY 481 CTTGGTGTTCCTATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTT 540
 Db 481 CTTGGTGTTCCTATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTT 540
 QY 541 CGCTATATCGGTGGAACAAACGCAACAGCTTCTCCAGTTCATGATGAACATCATCAAC 600
 Db 541 CGCTATATCGGTGGAACAAACGCAACAGCTTCTCCAGTTCATGATGAACATCATCAAC 600
 QY 601 GGTGGGCTCAGCTGACTCGGTGTTGACGTTTCCAGGCTTCCAGGCTTCCGATCGGT 660
 Db 601 GGTGGGCTCAGCTGACTCGGTGTTGACGTTTCCAGGCTTCCAGGCTTCCGATCGGT 660
 QY 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGTTACACGCACTGAAGTCC 720
 Db 661 GCAGAGACTTCTCTGAGGCTCTCCGCAACGGCGGAGGTTACACGCACTGAAGTCC 720
 QY 721 GTCATCAAGAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCTTCC 780
 Db 721 GTCATCAAGAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCTTCC 780
 QY 781 GTGGCTCAGCCGTGAGGCTCTGACCTTATGCTTGAAGCAATCGAAGGCTTGGCTTC 840
 Db 781 GTGGCTCAGCCGTGAGGCTCTGACCTTATGCTTGAAGCAATCGAAGGCTTGGCTTC 840
 QY 841 ACCCAGCAAGGACATCGCTTCTGCTCTGAGCTTCTCTCTGAGTCTTCTCAAGGAC 900
 Db 841 ACCCAGCAAGGACATCGCTTCTGCTCTGAGCTTCTCTCTGAGTCTTCTCAAGGAC 900
 QY 901 GGCACTTACCACTTCGAAGGTGGCGAGCTCCGAGCTGAGATGGCAAAAGCTTTACGCT 960
 Db 901 GGCACTTACCACTTCGAAGGTGGCGAGCTCCGAGCTGAGATGGCAAAAGCTTTACGCT 960
 QY 961 GAGCTCTGAGCGGTAACCAATCGTCTCATCGAGGACCCACTGCGAGGAAGATGACTGG 1020
 Db 961 GAGCTCTGAGCGGTAACCAATCGTCTCATCGAGGACCCACTGCGAGGAAGATGACTGG 1020
 QY 1021 GAGGTTTACCAACCTTCAAGCAACCATCGGCGAAGGTTTCAAGCTTGGCGAGC 1080
 Db 1021 GAGGTTTACCAACCTTCAAGCAACCATCGGCGAAGGTTTCAAGCTTGGCGAGC 1080
 QY 1081 TTCTTCTGCAACCACTTGAAGGCTTGAAGGAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 1140
 Db 1081 TTCTTCTGCAACCACTTGAAGGCTTGAAGGAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 1140
 QY 1141 ATCTGTTTGAAGTGAACCAAGATCGGTACCTTCAAGGACCTTTCGAGCTGTCGATG 1200
 Db 1141 ATCTGTTTGAAGTGAACCAAGATCGGTACCTTCAAGGACCTTTCGAGCTGTCGATG 1200
 QY 1201 GCTCAGCGGAGGCTTACCTTCAATGATGCTCCACCGTTCGAGGAGCAAC 1260
 Db 1201 GCTCAGCGGAGGCTTACCTTCAATGATGCTCCACCGTTCGAGGAGCAAC 1260
 QY 1261 ACCATTGCTGACCTCGGAGTTGCACTCACTGTGGCCAGATCAAGACTGTTGCTCCAGCA 1320

Db 1261 ACCATTGCTGACCTCGGAGTTGCACTCAACTGTGGCAGATCAAGACTGTTGCTCCAGCA 1320
 QY 1321 CGTTTCGACCGTGTTCGCAAAAGTACAACAGCTTCTCGCATCGAGAGCTGCTTGGCGAC 1380
 Db 1321 CGTTTCGACCGTGTTCGCAAAAGTACAACAGCTTCTCGCATCGAGAGCTGCTTGGCGAC 1380
 QY 1381 GCCGGGCTTACGAGGTCGAGCGCATTCGCCAGCTTTCCAGGCTTAAATAAAGCGCTT 1440
 Db 1381 GCCGGGCTTACGAGGTCGAGCGCATTCGCCAGCTTTCCAGGCTTAAATAAAGCGCTT 1440
 QY 1441 TTGACGCCCGGTAACTCAAGGTTGCCGGGCTGCTTGTCTTACTACTGTTACTGTTGT 1500
 Db 1441 TTGACGCCCGGTAACTCAAGGTTGCCGGGCTGCTTGTCTTACTACTGTTACTGTTGT 1500
 QY 1501 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 1501 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 QY 1561 CAAGCAGGGAACGTGCTT 1578
 Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 3
 BD014844
 LOCUS
 DEFINITION Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid.
 ACCESSION BD014844
 VERSION BD014844.1 GI:22555651
 KEYWORDS JP 2001161380-A/1.
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 (bases 1 to 1578)
 AUTHORS Meckel,V., Pfeifferle,W., Hermann,T., Puehler,A., Kalinovski,J. and Baate,B.
 TITLE Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid
 JOURNAL Patent: JP 2001161380-A 1 19-JUN-2001;
 DEGUSSA HUELS AG
 COMMENT OS Corynebacterium glutamicum
 PN JP 2001161380-A/1
 PD 19-JUN-2001
 PF 04-OCT-2000 JP 2000305110
 PR 05-OCT-1999 DE 19947791.4
 PI VETTYNA MECKEL, WALTER PFEIFFERLE, THOMAS HERMANN, ALFRED PUEHLER, JOERN KALINOVSKI, BRIGITTE BAATE
 PC C12N15/09, C12N1/21, C12N9/88, C12P13/04, C12P13/08, C12N15/09, C12R1:15),
 PC C12R1:15),
 PC (C12P13/04, C12R1:15), (C12P13/08, C12R1:15), C12N15/00, (C12N15/00, PC C12R1:15)
 CC Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid
 CC CC Polynucleotide, DNA, isolation of polynucleotide sequence, and method of fermentative production of L-amino acid
 PH Key Location/Qualifiers
 FT CDS Location/Qualifiers
 FEATURES
 source 1..1578
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 ORIGIN
 Query Match 100.0%; Score 1578; DB 6; Length 1578;
 Best Local Similarity 100.0%; Pred. No. 6.9e-288;
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCTGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTCGCTTCATCGAAACAAGA 60

Db	1	GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTGGCTTCATCGAAACAAGA	60
Qy	61	TTCTGTGCAAACAATTTGGGTGAGACGTGATTTGAAGACATTTTGATCACGTGAATAATTCTAG	120
Db	61	TTCTGTGCAAACAATTTGGGTGAGACGTGATTTGAAGACATTTTGATCACGTGAATAATTCTAG	120
Qy	121	TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCA CGTATTTCGTCTGC	180
Db	121	TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCA CGTATTTCGTCTGC	180
Qy	181	GAATTTCTCGACTCCCGCGGTAAACCBAACCGTTCGAGCGAGAGGTTTTCCTGGATGACGGT	240
Db	181	GAATTTCTCGACTCCCGCGGTAAACCBAACCGTTCGAGCGAGAGGTTTTCCTGGATGACGGT	240
Qy	241	TCCACACGGTGTCCGAGGTGTTTCCATTCGGCGGCATCCACCGCGTCCACGAGGCTCATGAG	300
Db	241	TCCACACGGTGTCCGAGGTGTTTCCATTCGGCGGCATCCACCGCGTCCACGAGGCTCATGAG	300
Qy	301	CTGCGTGACGGTGGCGATCGCTACTCTGGGCAAGGGCGTTTTTGAAGGCAGTTGAAAACGTC	360
Db	301	CTGCGTGACGGTGGCGATCGCTACTCTGGGCAAGGGCGTTTTTGAAGGCAGTTGAAAACGTC	360
Qy	361	AACGAAGAAATTCGGCGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCGCTCATCGAC	420
Db	361	AACGAAGAAATTCGGCGACGAGCTCGCTGGGCTTAGAGGCTGACGATCAGCGCTCATCGAC	420
Qy	421	GAAGCAATGATCAAGCTTGATGGCACCGCCAAACAAGTCCCGCTGGGTGCAACGCGAATC	480
Db	421	GAAGCAATGATCAAGCTTGATGGCACCGCCAAACAAGTCCCGCTGGGTGCAACGCGAATC	480
Qy	481	CTTGGTGTTCCTATGGCTGTTTGCAAAAGGCTGCTGATTCGGCAGGCTCCCACTGTTC	540
Db	481	CTTGGTGTTCCTATGGCTGTTTGCAAAAGGCTGCTGATTCGGCAGGCTCCCACTGTTC	540
Qy	541	CGCTACATTCGGTGGACAAAGCAACA CGTTCTTCCAGTTCCAATGATGAACATCATCAAC	600
Db	541	CGCTACATTCGGTGGACAAAGCAACA CGTTCTTCCAGTTCCAATGATGAACATCATCAAC	600
Qy	601	GGTGGCGCTCACGCTGACTCCGGTGTGACGTTTCAGAAATTCATGATCGCTCCAAATCGGT	660
Db	601	GGTGGCGCTCACGCTGACTCCGGTGTGACGTTTCAGAAATTCATGATCGCTCCAAATCGGT	660
Qy	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAA CGGCGCGGAGGTCTACCA CGCACTGAAGTCC	720
Db	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAA CGGCGCGGAGGTCTACCA CGCACTGAAGTCC	720
Qy	721	GTCAATCAAGGAATAAGGCGCTGTCAACCGGACTTGGCGATGAGGGCGGCTTCGTCTCTTCC	780
Db	721	GTCAATCAAGGAATAAGGCGCTGTCAACCGGACTTGGCGATGAGGGCGGCTTCGTCTCTTCC	780
Qy	781	GTCCGCTCCACCGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTC	840
Db	781	GTCCGCTCCACCGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTC	840
Qy	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGACGCTTCTTCCTCTGAGTTCTTCAAGGAC	900
Db	841	ACCCAGGCAAGGACATCGCTCTTGCTCTGACGCTTCTTCCTCTGAGTTCTTCAAGGAC	900
Qy	901	GGCACTTACCACTTCGAAAGGTGGCCAGCACTTCGCGAGCTGAGATGGCAAAACGTTTACGCT	960
Db	901	GGCACTTACCACTTCGAAAGGTGGCCAGCACTTCGCGAGCTGAGATGGCAAAACGTTTACGCT	960
Qy	961	GAGCTCGTTGACGGTACCCAAATCGTCTCCATCGAGGACCCCATGCGAGGAAGATGACTGG	1020
Db	961	GAGCTCGTTGACGGTACCCAAATCGTCTCCATCGAGGACCCCATGCGAGGAAGATGACTGG	1020
Qy	1021	GAGGGTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC	1080
Db	1021	GAGGGTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC	1080
Qy	1081	TTCTTTCGTCTACCAACCTCTGACGCGCTGAAGGAGGGCACTCGCTAAGAAAGGCTGCCAATCC	1140

Db	1081	TTCTTTCGTACCAACCCCTGAGCGCCCTGAAGAGGAGGCATCGCTAAGAAGGCTGCCAACTCC	1141
Qy	1141	ATCCTGGTTAAAGGTGAACACAGATCGGTACCTCCTCACCGAGACCTTCGACGCTGTCGACATG	1200
Db	1141	ATCCTGGTTAAAGGTGAACACAGATCGGTACCTCCTCACCGAGACCTTCGACGCTGTCGACATG	1200
Qy	1201	GCTCACCGCGCAGGGCTACACCTCCATGATGTCCTCCACCGTTCCGGTGAGACCGAGGACACC	1260
Db	1201	GCTCACCGCGCAGGGCTACACCTCCATGATGTCCTCCACCGTTCCGGTGAGACCGAGGACACC	1260
Qy	1261	ACCATTTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGCTGTCCTCAGCA	1320
Db	1261	ACCATTTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGCTGTCCTCAGCA	1320
Qy	1321	CGTTCCGACCGTGTGCGAAAGTACAACACAGCTTCTCCGATCGACGAGCTGCTTTGGCGAC	1380
Db	1321	CGTTCCGACCGTGTGCGAAAGTACAACACAGCTTCTCCGATCGACGAGCTGCTTTGGCGAC	1380
Qy	1381	GC CGGGCTGTACGACAGTTCGACGCGCATTTCCACGCTTTTCAGGGCTAAATAAAGCGCTT	1440
Db	1381	GC CGGGCTGTACGACAGTTCGACGCGCATTTCCACGCTTTTCAGGGCTAAATAAAGCGCTT	1440
Qy	1441	TTCGACGCCCGGTAAACCTCAAGTTGCGGGCGTCTGTTGCTTACTACTGTTACTGTTGT	1500
Db	1441	TTCGACGCCCGGTAAACCTCAAGTTGCGGGCGTCTGTTGCTTACTACTGTTACTGTTGT	1500
Qy	1501	GACTATGATCGAGGATTTATGGCAAGCAGAGAAGAACTCATAAAGGCTTTGTTCCCTGTCT	1560
Db	1501	GACTATGATCGAGGATTTATGGCAAGCAGAGAAGAACTCATAAAGGCTTTGTTCCCTGTCT	1560
Qy	1561	CAAGCAGGGAACGTGCTT 1578	
Db	1561	CAAGCAGGGAACGTGCTT 1578	
RESULT 4			
AP005277			
LOCUS			
DEFINITION	AP005277	333150 bp	DNA linear BCT 08-AUG-2002
ACCESSION	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 4/10		
VERSION	AP005277	BA000036	
KEYWORDS	AP005277.1	GI:21323710	
SOURCE	Corynebacterium glutamicum ATCC 13032		
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
REFERENCE	1 Nakagawa, S.		
AUTHORS	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 333150)		
REFERENCE	Nakagawa, S.		
AUTHORS	Direct Submission		
TITLE	Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa HAKKO Kogyo Co.		
JOURNAL	Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,		
	Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com,		
	Tel:81-44-829-3031, Fax:81-44-813-1651)		
	This sequence is conducted by collaboration of Kyowa HAKKO Kogyo		
COMMENT	Co. Ltd. And Kitasato University.		
FEATURES	Location/Qualifiers		
source	1. .333150		
	/organism="Corynebacterium glutamicum ATCC 13032"		
	/mol_type="genomic DNA"		
	/strain="ATCC 13032"		
	/db_xref="taxon:196627"		
	/note="ATCC 13032"		
gene	complement (63..1292)		
	/gene="Cgl10944"		
CDS	complement (63..1292)		
	/gene="Cgl10944"		
	/codon_start=1		
	/transl_table=11		

```

/product="Hypothetical membrane protein"
/db_xref="BAB98337.1"
/db_xref="GI:21323711"
/translacion="MTNPTERNARRLIWANGLQNIQDOIVAAKTVLPWLLQAAAGP
FLALLVPIRAGSMLOAAITGWLVQTSKVVIGSGQFVSALGIGVAAFLRAG
WALGITVIVLAAALSRSCSIASKOVQKVISKGRGLVTRATVIGVMGLVAGL
AIAIFLGSHTRVILAAVVIASSFSWLFASIVFAKIEYAKPATPKNAPSANPVRRCI
AALXDKAFRRFVLVRSMVLTAFTAPIVALAAESGNSIDSLGFFLIASGLASMVG
RISIGSDHSKNYMGAGLFGSIVLILVLSAPAPAQINLVFPLSFFLITLAHTA
IRVARKYVMDAEGDQRTYVADANTLMGVLLIVGALSFTIAPGNEALLFLAAI
GLLGTISARGLKEYSAG"
1382..2917
/gene="Cg10945"
1382..2917
/gene="Cg10945"
/codon_start=1
/transl_table=1
/product="Putative multicopper oxidases"
/protein_id="BAB98338.1"
/db_xref="GI:21323712"
/translacion="MPELNRRTPFKGAGVLAATVVGQVLVACSDDDVRVGYGSPRTL
PIPADLGTREGSVHFALBAOTGESQILPDVTTKWTGNGTHLGPTLVVKKGDDVHV
DVNNLDEMVTVMHGMKLPALADGPHSPIGPQOTWSPPTWVANDAATLWYHPHTG
LTGLHAYRGLAGLIIIVDEATDKLDPREYGVDDIPLVMDHRLFEGSDLEEDLPDL
GLGDDTPANGITWAHFDAITRRVRFRVNGNMRFYNLAFSDTRTFQVIAASDGLD
EPDRTTLAIGGERWIVVEFEGEDVTLESVGFEDNYGVPDDFDFDQMSDFOL
LITITPSDDAAAPALPGVLVKSTEPDVIDATERTFIMTFTSINDLQMDQMDVDVID
HDQEVMIVTNDSNDWPHNFVHDAREFKVLKFECDVLEFNDGKQTVGLPGCATL
AVEGHYDPQWPYMYCHMLYHEDQGMQGVIVFPGDEPAAVLGSCTGSSIDSAGG
HAH"
3063..3983
/gene="Cg10946"
3063..3983
/gene="Cg10946"
/notes="PF00005:ABC transporter
TIGR00960:3a0501802: Type II (General) Secretory Pathway
(IISF) Family protein
TIGR00968:3a0106801: sulfate transport system permease
protein
TIGR01166:cbio: cobalt transport protein ATP-binding
subunit
TIGR01184:ntrCD: nitrate transport ATP-binding subunits C
and D
TIGR01186:prov: glycine betaine/L-proline transport ATP
binding subunit
TIGR01187:potA: spermidine/putrescine ABC transporter
ATP-binding subunit
TIGR01188:drda: daunorubicin resistance ABC transporter
ATP-binding subunit
TIGR01189:ccmA: heme exporter protein Ccma"
/codon_start=1
/transl_table=1
/product="ABC-type transporter, ATPase component"
/protein_id="BAB98339.1"
/db_xref="GI:21323713"
/translacion="MTTTPAIDVTLVRYTGYDTAVKGLNFHVHVGFLIGTNGAG
KTSLEVIEGLSAPSSCTVIRISGLDPVADRAILRPELIMLQSGLPSPQVTAETMDM
WHTCTYPAIKOVLDVADLHRENVKVGALSQGEORRLACALIGDPSILFDEPT
GLDPERRHTWQLLDLQKRGVMTLTHYLEAEFLCDRIAIMNAGEIAVGTGLDE
LVAREKSIISFVLRGQGVFLPSGAEIIRDNNHVRTATTLOOHTLEILTAAETGI
ALGCFAPAKPATLESVFMDIASLENTSLQTA"
4003..4836
/gene="Cg10947"
4003..4836
/gene="Cg10947"
/codon_start=1
/transl_table=1
/product="ABC-type transporter, permease components"
/protein_id="BAB98340.1"
/db_xref="GI:21323714"
/translacion="MTTSHTARGLOHATPERRKTSFFKTSLFKAEMIQFRNKTLFLM
ATVPVGIPLLLIINGGAASANSFDYFVMVTLFLVQFTVLSMATRRDRLVKR
LRTGEARDIDIIGAICPGALLTLFTWIIPLMLVLAGAPINLVIFVAVLIGLL

```

```

CSALALMTSGFTRNAEAQNTSMFPVFLMAMGLGSIIRFVFGDSIVADILAYTPFAAIS
DLVQIGWAGATFADSVGVEAANFAGIFQDMLIPLGILAAWTAATAAANRMYRWSY
R"
4879..6138
/gene="Cg10948"
4879..6138
/gene="Cg10948"
/EC number="2.7.3.-"
/codon_start=1
/transl_table=1
/product="Two-component system, sensory transduction
histidine kinases"
/protein_id="BAB98341.1"
/db_xref="GI:21323715"
/translacion="MNKDFWTAGTWARTWFRSGVSLASPTVPLNSWRRRLPNLAKYTL
YTRSLQAIPIVLLSAYFLGIVANAGTLNPSFVWLGLGFSVILLIIVLVVYQPSLNS
HPRSVQPPFFTGLVLNVGVVSVVLIQIPLGNNSDNTRATALIFTLTCTVLLSIAYI
PMNRYRWLIAMSVLWMTSTTIDYLSALWVIPPPLMAGTVRLSVTVDMKVEVRS
RELEASLRVTEERLRFQAEHLDTLGOHAAWSKSELALAKRGDDRLNELELQK
LRTTSMSEMDVSVGYRTVNLAITEIGAKSILLADAHLISLVIGTTSQVSPAHRELCAW
LVREATTNLRHSDATDATLTLSSTEVMDNNGVNDIGRLSLGSLARSRAESAGMTL
IVSREDDQFSVRMLINAPANTPAEKA"
6139..6750
/gene="Cg10949"
6139..6750
/gene="Cg10949"
/notes="PF00072:Response regulator receiver domain
PF00196:Bacterial regulatory proteins, luxR family"
/codon_start=1
/transl_table=1
/product="Two-component system, response regulators
consisting of a CheY-like receiver domain and a HTH
DNA-binding domain"
/protein_id="BAB98342.1"
/db_xref="GI:21323716"
/translacion="MISISIADEALTIASSLATLLSLEPDLDPRTAGSGEELIETWA
DPNRTDVCVLDIQLGIDIGIDITRLMETTPDLAVITVSHARPROUKRALAAGVLG
FLPKTSTADEFATAIRTVHAGRYIDPELAAMTISAGESPLTNEEVEVLELAGQGSLA
EETAAVAHLAPGTTNVLISOAMTKVGAQRFEAFTRARELGWL"
complement(6781..6987)
/gene="Cg10950"
complement(6781..6987)
/gene="Cg10950"
/codon_start=1
/transl_table=1
/product="Hypothetical protein"
/protein_id="BAB98343.1"
/db_xref="GI:21323717"
/translacion="MKHYMSMYFGAGPLNTPHTKEVMFMDIVSIVNHFWEFSTNNWVN
MTGLPRETFVPAFVGVSTVLAVLSF"
7048..8205
/gene="Cg10951"
7048..8205
/gene="Cg10951"
/codon_start=1
/transl_table=1
/product="rRNA-processing ribonuclease BN"
/protein_id="BAB98344.1"
/db_xref="GI:21323718"
/translacion="MTAWLCSPFSPQVDEYDMNGVVPQOEHLDATLIAADFHGPNP
SGRKLRLNPFQWKYALNRTVRDPDGLDLAALLTFPSILSTAPAVLLGYSVITIP
LASDSTEILNLVDEYVNOYVPEDOSHVVNGVIDSIAGSAAGOVGAVGVITALTSS
AYVAFSRCANAVYGRSEKTLTKRWMLFLNLALLGLIIILVSVWLNELTMGFI
APIAEPLHNLTVLSFLITDRPMPIWIWVFPVIVGVLMFVATLYWAPNARPMKFRWL
SLGSFLAIVGILLAGLNFYFTLFAAFSSYGAVGSLAVFIALWVFNICLIITGLKID
VEISRAKQLQAGMGAEDYSLVPPRSIEKVAKMQRQORLMDQAAARESN"
8672..10114
/gene="Cg10952"
8672..10114
/gene="Cg10952"
/notes="PF00005:ABC transporter
PF00664:ABC transporter transmembrane region."
/codon_start=1

```

```
/transl_table=11
/product="ABC-type multidrug/protein/lipid transport
system, AtPase component"
/protein_id="BAB98345.1"
/db_xref="GI:21323719"
/translation="MRRTHTAGSILLRMRORGVAFGAFPLGMWOLSEALVPLAI
GLIVDHAVLTKDLRLVGLVAPVVLFSYRFGSRALNRVNFESHARVEVAD
HALKNDLPRNLVGEVMSRSTADSDSTIFGQIGTGVSAAATFLGAATYLLISDMLV
GLIVLVLVPIIISGVALVASKGI SKRSTVQOEKLAESQAASDIMGLRVYKAIGGSRW
AVTFEKASQASAPRIILSNSINASAIHAAGAEVANFLNLAALVLLAGWRVYTGELGPQ
LIAIVGAVLSPRIILSNSINASAIHAAGAEVANFLNLAALVLLAGWRVYTGELGPQ
FLIVPPASTPLPHGDNILATPHAADIFEGTLRSNI SMNHEDNVPIDPOVIRASGLTDI
IEVDGLDAPVRDTSNLSGGQRQVALARALHADAEVLVLMDFTSADVSTVEVSIAGG
IKQRRAGCTIVVSSSPAFYNLADRVISHV"
gene 10107. .11843
CDS /gene="Cg10953"
10107. .11843
/gene="Cg10953"

Query Match 100.0%; Score 1578; DB 1; Length 333150;
Best Local Similarity 100.0%; Pred. No. 4,9e-288;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTGGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTGCCCTCATCGAAACAAGA 60
Db 34849 GGCTGGGGATATGGTAGTTTTCGCCCACTAATTTCAACTGATTGCCCTCATCGAAACAAGA 34908

Qy 61 TTCTGTCAACAATTTGGGTAGACGTGATTTGAAGACATTTTGAACGTGAATTAATCTAG 120
Db 34909 TTCTGTCAACAATTTGGGTAGACGTGATTTGAAGACATTTGATCACTGTAATTAATCTAG 34968

Qy 121 TTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAATCATGCAAGTATTGCCTCGC 180
Db 34969 TTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAATCATGCAAGTATTGCCTCGC 35028

Qy 181 GAATTTCTCGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGACGGT 240
Db 35029 GAATTTCTCGACTCCCGCGGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGACGGT 35088

Qy 241 TCCACAGGTGTCGAGGTGTTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 300
Db 35089 TCCACAGGTGTCGAGGTGTTTCCATCCGGCGCATCCACCGCGCTCCACGAGGCTCATGAG 35148

Qy 301 CTGGGTGACGGTGGCGATCGTACTCTGGGCAAGGGCGTTTGAAGGACAGTTGAACAGTC 360
Db 35149 CTGGGTGACGGTGGCGATCGTACTCTGGGCAAGGGCGTTTGAAGGACAGTTGAACAGTC 35208

Qy 361 AACGAAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
Db 35209 AACGAAGAAATCGGCGACGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 35268

Qy 421 GAAGCAATGATCAAGCTTGATGGACCGCCCAACAAAGTCCCGCTGGGTGCAACGCAATC 480
Db 35269 GAAGCAATGATCAAGCTTGATGGACCGCCCAACAAAGTCCCGCTGGGTGCAACGCAATC 35328

Qy 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
Db 35329 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 35388

Qy 541 CGCTACATCGGTGACCAACGCAACAGTTCCTCCAGTTTCCAAATGATGAACATCATCAAC 600
Db 35389 CGCTACATCGGTGACCAACGCAACAGTTCCTCCAGTTTCCAAATGATGAACATCATCAAC 35448

Qy 601 GGTGGCGCTACGCTGACTCGCGTGTGACGTTTCAGGAATTCATGATCGCTCCCAATCGGT 660
Db 35449 GGTGGCGCTACGCTGACTCGCGTGTGACGTTTCAGGAATTCATGATCGCTCCCAATCGGT 35508

Qy 661 GCAGAGACTTCTCTGAGGCTCTCGGCAACGGCGGAGGCTTACCACGCACTGAAGTCC 720
Db 35509 GCAGAGACTTCTCTGAGGCTCTCGGCAACGGCGGAGGCTTACCACGCACTGAAGTCC 35568

Qy 721 GTCATCAAGGAAAAGGGCTGTGCAACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
Db 35569 GTCATCAAGGAAAAGGGCTGTGCAACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 35628

Qy 781 GTCGGCTCAACCGGTGAGGCTCTTGATCTGTTAGGCAATCGAGAAAGGCTTGGCTTC 840
Db 35629 GTCGGCTCAACCGGTGAGGCTCTTGATCTGTTAGGCAATCGAGAAAGGCTTGGCTTC 35688

Qy 841 ACCCAGGCAAGACATCGCTCTTGCTCTGGAGTTTCTTCCCTGAGTTCTTCAAGGAC 900
Db 35689 ACCCAGGCAAGACATCGCTCTTGCTCTGGAGTTTCTTCCCTGAGTTCTTCAAGGAC 35748

Qy 901 GGCACCTACCACTTCGAAGTGCCAGCACTCCGCAAGCTGAGATGGCAACGTTTACGCT 960
Db 35749 GGCACCTACCACTTCGAAGTGCCAGCACTCCGCAAGCTGAGATGGCAACGTTTACGCT 35808

Qy 961 GAGCTCGTTGACCGGTACCCCAATCGTCTCCATCGAGACCCACTGCGAGGAAGATGACTGG 1020
Db 35809 GAGCTCGTTGACCGGTACCCCAATCGTCTCCATCGAGACCCACTGCGAGGAAGATGACTGG 35868

Qy 1021 GAGGCTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTCCAGATCGTTGGCGACGAC 1080
Db 35869 GAGGCTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTCCAGATCGTTGGCGACGAC 35928

Qy 1081 TTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCGCAACTCC 1140
Db 35929 TTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCGCAACTCC 35988

Qy 1141 ATCTCTGTTAAGTGAAACAGATCGGTACCTCACCGAGACCTTTCGACGCTGTCGACATG 1200
Db 35989 ATCTCTGTTAAGTGAAACAGATCGGTACCTCACCGAGACCTTTCGACGCTGTCGACATG 36048

Qy 1201 GCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCGGAGCACCC 1260
Db 36049 GCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCGGAGCACCC 36108

Qy 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGGCGACATCAAGACTGCTGCTCCAGCA 1320
Db 36109 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGGCGACATCAAGACTGCTGCTCCAGCA 36168

Qy 1321 CGTTTCGACCGGTGCGAAAGTACAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Db 36169 CGTTTCGACCGGTGCGAAAGTACAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 36228

Qy 1381 GCCGGCGTCTACCGAGGTGCGAGCGATTCGCCAGCTTTCAGGCTTAAATAAAGCGCTT 1440
Db 36229 GCCGGCGTCTACCGAGGTGCGAGCGATTCGCCAGCTTTCAGGCTTAAATAAAGCGCTT 36288

Qy 1441 TTCGACGCCCGGTAAACCTCAAGGTTGCGGGCGTCTGCTTACTCTGTTACTGCTGT 1500
Db 36289 TTCGACGCCCGGTAAACCTCAAGGTTGCGGGCGTCTGCTTACTCTGTTACTGCTGT 36348

Qy 1501 GACTATGATCGAGGATTTATGGCAAGAGCAAGAAAACCTCAAGGCTTGTTCCTGCTCT 1560
Db 36349 GACTATGATCGAGGATTTATGGCAAGAGCAAGAAAACCTCAAGGCTTGTTCCTGCTCT 36408

Qy 1561 CAAGCAGGAACGTCGCTT 1578
Db 36409 CAAGCAGGAACGTCGCTT 36426

RESULT 5
BX927150
LOCUS BX927150 348475 bp DNA linear BCT 10-JUN-2004
DEFINITION Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5,
complete genome; segment 3/10.
ACCESSION BX927150 BX927147
VERSION BX927150.1 GI:41324904
KEYWORDS complete genome.
SOURCE Corynebacterium glutamicum ATCC 13032
ORGANISM Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 348475)
AUTHORS Kalinowski,J., Bathe,B., Bartels,D., Biscoff,N., Bott,M.,
```


gene	7020. .8429	Db	338771	AACGAAGAAATCGCGACGAGCTCGCTGGCTAGAGGCTAGACGATCAGCGCCTCATCGAC	338830
CDS	/gene="lpda"	QY	421	GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTGGGTGCAAAAGCGCAATC	480
	7020. .8429	Db	338831	GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTGGGTGCAAAAGCGCAATC	338890
	/gene="lpda"	QY	481	CTTGGTGTTCATGCTGTGTCAAAGGCTGCTGATTCGCGAGGCTCCCACTGCTTC	540
	/locus_tag="cg0790"	Db	338891	CTTGGTGTTCATGCTGTGTCAAAGGCTGCTGATTCGCGAGGCTCCCACTGCTTC	338950
	/EC_number="1.8.1.4"	QY	541	CGCTACATCGGTGCGACAAACGACACGCTTCTCCAGTTCCTCAATGATGAACATCATCAAC	600
	/codon_start=1	Db	338951	CGCTACATCGGTGCGACAAACGACACGCTTCTCCAGTTCCTCAATGATGAACATCATCAAC	339010
	/transl_table=11	QY	601	GGTGGCGCTCACGCTGACTCCGCTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT	660
	/product="DIHYDROLIPOAMIDE DEHYDROGENASE"	Db	339011	GGTGGCGCTCACGCTGACTCCGCTGTTGAGCTTCAGGAATTCATGATCGCTCCCAATCGGT	339070
	/protein_id="CAF19393.1"	QY	661	GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACACGACCTGAAGTCC	720
	/db_xref="GI:41324912"	Db	339071	GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGGCTTACACGACCTGAAGTCC	339130
	/translations="MAKRIVIIGGPGAGYEAALAGAKYGAETVVIDVGVSATVMD	QY	721	GTCTCAAGGAAAGGCGCTTCTCCACCGACTTTGGCGATGAGGCGCGCTTCGCTCCTCC	780
	CVPKSFIACTGKTLRRADDMGLNRGLGKHALEIDALNIRVKDLAKAQSDILGOL	Db	339131	GTCTCAAGGAAAGGCGCTTCTCCACCGACTTTGGCGATGAGGCGCGCTTCGCTCCTCC	339190
	QRSQVMIINGVGFEDYNTKQTHYIKVTHSDSESTVECDLVLVATGATPRLLKGA	QY	781	GTCCGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC	840
	PDGERLLTRQVTDIELPHTLLVVGSGVTFSAFAELGVKVTWASRDRLPHD	Db	339191	GTCCGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTC	339250
	DADAQVLETVAERGSPLEKHARVESVTRTDDGVCVTRDAGDCTDLFFPLASVAAQMGRTAM	QY	841	ACCCAGGCAAGGACATCGCTCTTGCTCGAGGCTTCTTCCTCTGAGTTCTTCAAGGAC	900
	PNTADLGNIGVAPSGHIKVDVRSRTNIPGVYAAGDCTDLFFPLASVAAQMGRTAM	Db	339251	ACCCAGGCAAGGACATCGCTCTTGCTCGAGGCTTCTTCCTCTGAGTTCTTCAAGGAC	339310
	YHALGVSPIRLUKTAVATVTRPEIAAAGVITHAQVDSGVSARVTVLPLATNPRAM	QY	901	GGCACCTTACACTTCGAAGGTGCCAGCACTCCCGAGCTGAGATGGCAAAAGCTTACGCT	960
	YSLRHGFVKLFRRNSGLIIGGVVAPTASELILPIAVAVTNRLTVADLADTFAVYPS	Db	339311	GGCACCTTACACTTCGAAGGTGCCAGCACTCCCGAGCTGAGATGGCAAAAGCTTACGCT	339370
	LSGSITENARQLVQDDLG"	QY	961	GAGCTGTTGACGGGTACCAATCGTCTCCATCGAGGACCACTGCGAGGAGTACTGCG	1020
	8826. .12248	Db	339371	GAGCTGTTGACGGGTACCAATCGTCTCCATCGAGGACCACTGCGAGGAGTACTGCG	339430
	/gene="pyc"	QY	1021	GAGGCTTACACCAACCTTCCCGCAACCATCGCGGCAAGGTTTCAGATCGTTGGCGAGAC	1080
	/locus_tag="cg0791"	Db	339431	GAGGCTTACACCAACCTTCCCGCAACCATCGCGGCAAGGTTTCAGATCGTTGGCGAGAC	339490
	8826. .12248	QY	1081	TTCTTTCGTACCAACCTTCCCGGCTGAGCGCTTGAAGGAGGATCGCTAAGAGGCTGCCAATCC	1140
	/gene="pyc"	Db	339491	TTCTTTCGTACCAACCTTCCCGGCTGAGCGCTTGAAGGAGGATCGCTAAGAGGCTGCCAATCC	339550
	/locus_tag="cg0791"	QY	1141	ATCCTGGTTAAGGTGAACAGATCGGTACCTTCCACGAGACCTTCGACGCTGTCGACATG	1200
	/EC_number="6.4.1.1"	Db	339551	ATCCTGGTTAAGGTGAACAGATCGGTACCTTCCACGAGACCTTCGACGCTGTCGACATG	339610
	/codon_start=1	QY	1201	GCTTACCGCGCAGGCTACACTCCATGATTCGACCTCCCGTTCGCTGAGACCGAGGACACC	1260
	/transl_table=11	Db	339611	GCTTACCGCGCAGGCTACACTCCATGATTCGACCTCCCGTTCGCTGAGACCGAGGACACC	339670
	/product="DIHYDROLIPOAMIDE DEHYDROGENASE"	QY	1261	ACCATTCGTGACCTTCGAGTTGACCTCAACTGTGTGGCCAGATCAAGAATCGTGTCTCAGCA	1320
	/protein_id="CAF19393.1"	Db	339671	ACCATTCGTGACCTTCGAGTTGACCTCAACTGTGTGGCCAGATCAAGAATCGTGTCTCAGCA	339730
	/db_xref="GI:41324912"	QY	1321	CGTTCCGACCGCTGCGCAAAAGTACAACAGCTTCTCCGATTCGAGGAGCTGCTTGGCGAC	1380
	/translations="MAKRIVIIGGPGAGYEAALAGAKYGAETVVIDVGVSATVMD	Db	339731	CGTTCCGACCGCTGCGCAAAAGTACAACAGCTTCTCCGATTCGAGGAGCTGCTTGGCGAC	339790
	CVPKSFIACTGKTLRRADDMGLNRGLGKHALEIDALNIRVKDLAKAQSDILGOL	QY	1381	GCCGGGCTCTACGAGGTCGACGATTCCTCCAGCTTTTCAGGCTTAAATAAAAGCGCTT	1440
	QRSQVMIINGVGFEDYNTKQTHYIKVTHSDSESTVECDLVLVATGATPRLLKGA	Db	339791	GCCGGGCTCTACGAGGTCGACGATTCCTCCAGCTTTTCAGGCTTAAATAAAAGCGCTT	339850
	PDGERLLTRQVTDIELPHTLLVVGSGVTFSAFAELGVKVTWASRDRLPHD	QY	1441	TTTGAAGCGCCGGTAACTTCAAGGTTGCCGGGCTGCTGTGCTTACTACTGTACTGTGT	1500
	DADAQVLETVAERGSPLEKHARVESVTRTDDGVCVTRDAGDCTDLFFPLASVAAQMGRTAM	Db	339851	TTTGAAGCGCCGGTAACTTCAAGGTTGCCGGGCTGCTGTGCTTACTACTGTACTGTGT	339910
	PNTADLGNIGVAPSGHIKVDVRSRTNIPGVYAAGDCTDLFFPLASVAAQMGRTAM				
	YHALGVSPIRLUKTAVATVTRPEIAAAGVITHAQVDSGVSARVTVLPLATNPRAM				
	YSLRHGFVKLFRRNSGLIIGGVVAPTASELILPIAVAVTNRLTVADLADTFAVYPS				
	LSGSITENARQLVQDDLG"				
	8826. .12248				
	/gene="pyc"				
	/locus_tag="cg0791"				
	8826. .12248				
	/gene="pyc"				
	/locus_tag="cg0791"				
	/EC_number="6.4.1.1"				
	/codon_start=1				
	/transl_table=11				
	/product="PYRUVATE CARBOXYLASE"				
	/protein_id="CAF19394.1"				
	/db_xref="GI:41324913"				
	/translations="MSHTSTSLPAFKKILVANRGEIAVRAALLETGAATVAIYPR				
	EDRGSFHRSPASAVRIGTSGSPVKAYLDIDEIIGAUKYKADAIYPGYFLSENAQL				
	ARECANRGTIFGTEPVLDTGSKRAVTAAKKAGLPVLAESTPSKNIIDEIVKSAEG				
	QTPYIFVAVAGGGRMRVAFSGDELKLAETASREAAAFGDAVYVERAVINPOH				
	IEVOILGDHTEVVEVHYERDCSLQRHKKVEIAPAHLDPELDRICADAVKPCRSI				
	GYOGATVEPLVDKGNHVFEMNPRLOVEHTTVEVTEVDLKAOMRLAAGATLXEL				
	GLTQDKIKTHGAALQCRITTEDPNNRPTDGIITAYRSPGGAGVRLDGAQLGSGIT				
	AHFDMLVMTGCSDFETAARAKALAEFTVSGVATNIGFRLALREDEDFTSKRIA				
	TGFIADPHLLQAPPADEQGRILDVLADVTVNKPHGPVRAAIDKLPNIKLDLPL				
	Query Match				
	Best Local Similarity				
	Matches 1578; Conservative				
	100.0%; Score 1578; DB 1; Length 348475;				
	100.0%; Pred. No. 4,9e-288;				
	Matches 1578; Conservative				
	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	GGCTGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA	60		
Db	338411	GGCTGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA	338470		
QY	61	TTCTGTCACAAATTTGGGTAGACGTGATTAAGACATTTTGCATCGTGAATATTCCTAG	120		
Db	338471	TTCTGTCACAAATTTGGGTAGACGTGATTAAGACATTTTGCATCGTGAATATTCCTAG	338530		
QY	121	TTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTCAAAATCATGACGATTCGCTCGC	180		
Db	338531	TTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTCAAAATCATGACGATTCGCTCGC	338590		
QY	181	GAAATTCCTGACCTCCCGCGTAAACCAACCGTTCGAGGACAGAGTTTTCCTGGATGACGGT	240		
Db	338591	GAAATTCCTGACCTCCCGCGTAAACCAACCGTTCGAGGACAGAGTTTTCCTGGATGACGGT	338650		
QY	241	TCCACGGTTCGACAGGTTCATTCGGCGCATCCACCGGCTCCACGAGGCTCATGAG	300		
Db	338651	TCCACGGTTCGACAGGTTCATTCGGCGCATCCACCGGCTCCACGAGGCTCATGAG	338710		
QY	301	CTCGTGACGGTTCGACATCGTACCTCGGCAAGGCGCTTTTGAAGGAGTGTGAAGACGTC	360		
Db	338711	CTCGTGACGGTTCGACATCGTACCTCGGCAAGGCGCTTTTGAAGGAGTGTGAAGACGTC	338770		
QY	361	AACGAAGAAATCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC	420		

QY 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAAACCTCATAAAGGCTTGTTCCTGTCT 1560
|||||
Db 339911 GACTATGATCGAGGATTTATGGCAAGCAGAGAAAACCTCATAAAGGCTTGTTCCTGTCT 339970
|||||
QY 1561 CAAGCAGGAAACGTGCTT 1578
|||||
Db 339971 CAAGCAGGAAACGTGCTT 339988
|||||
RESULT 6
AX127146 AX127146 349980 bp DNA linear (PATENT#2001
LOCUS DEFECTION) Sequence 7062 from Patent EP1108790.
ACCESSION AX127146 AX114121
VERSION AX127146.1 GI:14041134
KEYWORDS Corynebacterium glutamicum
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE 1
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7062 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES Location/Qualifiers
source 1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3.309.400) split in 11, seq 7062
0.900.001 1.249.980"

Query Match 100.0%; Score 1578; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 4.9e-288;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATTTGCCTCATCGAAACAAGA 60
Db 134799 GGCTGGGGATATGGGTAGTTTTCGCCCACTAATTTCAACTGATTTGCCTCATCGAAACAAGA 134858
|||||
QY 61 TTCTGTCAACAATTTGGGTGTAGAGCTGATTGAAGACATTTGATCAGCTGAATATTCTAG 120
Db 134859 TTCTGTCAACAATTTGGGTGTAGAGCTGATTGAAGACATTTGATCAGCTGAATATTCTAG 134918
|||||
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTAATTCGCTCGC 180
Db 134919 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTAATTCGCTCGC 134978
|||||
QY 181 GAAATTCGACTCCCGGGTAACCCACCGTCAGGAGAGGTTTTCCTGGATGACGGT 240
Db 134979 GAAATTCGACTCCCGGGTAACCCACCGTCAGGAGAGGTTTTCCTGGATGACGGT 135038
|||||
QY 241 TCCACACGCTGTCAGAGTTTCCATCCGGGCGATCCACCGGCGTCCAGGGCTCATGAG 300
Db 135039 TCCACACGCTGTCAGAGTTTCCATCCGGGCGATCCACCGGCGTCCAGGGCTCATGAG 135098
|||||
QY 301 CTGCGTGACCGGTGGCGATCGCTACTCTGGCAAGGGCGTTTTTGAAGGAGGATTTGAACAGTC 360
Db 135099 CTGCGTGACCGGTGGCGATCGCTACTCTGGCAAGGGCGTTTTTGAAGGAGGATTTGAACAGTC 135158
|||||
QY 361 AACGAAGAAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACATCAGCGCCTCATCGAC 420
Db 135159 AACGAAGAAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACATCAGCGCCTCATCGAC 135218
|||||
QY 421 GAACCAATGATCAAGCTTGATGGCACCACCAAGTCCCGCTGGGTGCAACCGCAATC 480
Db 135219 GAACCAATGATCAAGCTTGATGGCACCACCAAGTCCCGCTGGGTGCAACCGCAATC 135278
|||||
QY 481 CTTGGTGTTCATGGCTGTTTGCAAAAGCTGCTGATTCGCGAGGCTTCCCACTGTTC 540
|||||

Db 135279 CTTGGTGTTCATGGCTGTTGCAAAAGGCTGCTGCTGATTTCCGAGGCTCCCACTGTTC 135338
|||||
QY 541 CGCTACATCGGTGAGCAACACGACAGTTCCTCCAGTCCCAATGATGAACATCATCAAC 600
|||||
Db 135339 CGCTACATCGGTGAGCAACACGACAGTTCCTCCAGTCCCAATGATGAACATCATCAAC 135398
|||||
QY 601 GGTGGGGCTCACGCTGACTCCGGTGTGTGAGCTTCAGGAAATTCATGATCGCTCCAAATCGGT 660
|||||
Db 135399 GGTGGGGCTCACGCTGACTCCGGTGTGTGAGCTTCAGGAAATTCATGATCGCTCCAAATCGGT 135458
|||||
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACACGACGACTGAAAGTCC 720
|||||
Db 135459 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTTACACGACGACTGAAAGTCC 135518
|||||
QY 721 GTCATCAAGGAAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
|||||
Db 135519 GTCATCAAGGAAAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 135578
|||||
QY 781 GTCGGCTTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAGAAAGGCTTGGCTTC 840
|||||
Db 135579 GTCGGCTTCCACCGCTGAGGCTCTTGACCTTATCGTTGAGGCAATTCGAGAAAGGCTTGGCTTC 135638
|||||
QY 841 ACCCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTCTGAGTTCCTTCAAGGAC 900
|||||
Db 135639 ACCCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTCTGAGTTCCTTCAAGGAC 135698
|||||
QY 901 GGCACCTTACCACCTTCGAAGGTGGCGACACTCCGCGAGCTCAGATGGCAAGGTTTACGCT 960
|||||
Db 135699 GGCACCTTACCACCTTCGAAGGTGGCGACACTCCGCGAGCTCAGATGGCAAGGTTTACGCT 135758
|||||
QY 961 GAGTCTGTTGACGCGTACCCCAATCGTCTCCATCGAGACCCACTTCGAGGAAGATGACTGG 1020
|||||
Db 135759 GAGTCTGTTGACGCGTACCCCAATCGTCTCCATCGAGGACCCACTTCGAGGAAGATGACTGG 135818
|||||
QY 1021 GAGGTTTACACCAACTTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC 1080
|||||
Db 135819 GAGGTTTACACCAACTTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGAC 135878
|||||
QY 1081 TTCTTGTGTACCAACCTCGAGCGCTGAAGGAGGCAATCGCTAAGAAGGCTGCCCACTCC 1140
|||||
Db 135879 TTCTTGTGTACCAACCTCGAGCGCTGAAGGAGGCAATCGCTAAGAAGGCTGCCCACTCC 135938
|||||
QY 1141 ATCTTGTGTTAAGGTGAACAGATCGGTACCTTCAACGAGACCTTCGACGCTGTCGACATG 1200
|||||
Db 135939 ATCTTGTGTTAAGGTGAACAGATCGGTACCTTCAACGAGACCTTCGACGCTGTCGACATG 135998
|||||
QY 1201 GCTCACGCGCAGGCTACACCTCCATGATGTCCCAACCGTTCGCGTGAGACCGAGGACACC 1260
|||||
Db 135999 GCTCACGCGCAGGCTACACCTCCATGATGTCCCAACCGTTCGCGTGAGACCGAGGACACC 136058
|||||
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTTGCTCCAGCA 1320
|||||
Db 136059 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTTGCTCCAGCA 136118
|||||
QY 1321 CGTTCCAGCAGCTGTCGCAAAAGTACAACAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGAC 1380
|||||
Db 136119 CGTTCCAGCAGCTGTCGCAAAAGTACAACAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGAC 136178
|||||
QY 1381 GCGCGGCTCTACGAGGTCGAGCGCATTTCCACGCTTTTCAGGGCTAAATAAAGCGGCTT 1440
|||||
Db 136179 GCGCGGCTCTACGAGGTCGAGCGCATTTCCACGCTTTTCAGGGCTAAATAAAGCGGCTT 136238
|||||
QY 1441 TTCCAGCGCCGGTAACTCAAGGTTGCGGGCGCTGTTGCTTACTGCTTACTGCTGT 1500
|||||
Db 136239 TTCCAGCGCCGGTAACTCAAGGTTGCGGGCGCTGTTGCTTACTGCTTACTGCTGT 136298
|||||
QY 1501 GACTATGATCGAGGATTTATGGCAAGCAGAGAAAACCTCATAAAGGCTTGTTCCTGTCT 1560
|||||
Db 136299 GACTATGATCGAGGATTTATGGCAAGCAGAGAAAACCTCATAAAGGCTTGTTCCTGTCT 136358
|||||
QY 1561 CAAGCAGGAAACGTGCTT 1578
|||||

[illegible]

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

1 Zelder, O., Pompejus, M., Schroeder, H., Kroeger, B., Klopffrogge, C. and
Habberhauer, G.

Genes encoding for carbon metabolism and energy-producing proteins
Patent: WO 03040291-A 7 15-MAY-2003;
BASP AKTIENGESSELLSCHAFT (DE)

FEATURES
Location/Qualifiers

1. 1405

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1718"

101. .1378

/note="unnamed protein product; RXA00235"

/codon_start=1

/transl_table=11

/protein_id="CAD99859.1"

/db_xref="GI:32257759"

/translation="MAEIMHVFAREIILDSRGNPTVEAEVFLDDSHGVAGVPSGASTG
VHEAHLRDGDRYLGKVLKAVENVEIGDELAEADDORLIIDAMIKLDGTANK
SRIGANAILGVSMVAKAAADSAGLPFRVIGGPNHVLVPMWNIITGGHADSGVD
VOEPMIPIAGETSEALRNGAEYHAKSVIKKGLSTGLDEGGFAPSVGSTREAL
DLVLEAKAGFTPGKIDIALALDVASSEFPKDYHPEGGHSAENVAELVDAY
PIVSIEDPLEDDWEGYTNLTATIGDKVQIVIGDDFFVTNPERLKEGIAKXAANSILVK
VNOIGTUTETFDVMAHRAGYTSMMSHRSGETEDTTIADIAVALNCCQIKTGAPARS
DRVAKYNQLLRILEQLLDAGVYAGRSAPPRFG"

ORIGIN

Query Match 88.9%; Score 1403.4; DB 6; Length 1405;

Best Local Similarity 99.9%; Pred. No. 6.6e-255;

Matches 1404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 51 CGAACAAAGATTCGTGCAACAATTTGGGTGTAGAGCTGATTGAAGACATTTGATCACGTGA 110

Db 1 CGAACAAAGATTCGTGCAACAATTTGGGTGTAGAGCTGATTGAAGACATTTGATCACGTGA 60

Qy 111 ATAAATTCAGTTAGTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAATCATGACAGT 170

Db 61 ATAAATTCAGTTAGTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAATCATGACAGT 120

Qy 171 ATTCGCTCGGAAATTTCTCGACTCCCGCGTAAACCCAAACCGTCGAGGACAGGTTTTCCT 230

Db 121 ATTCGCTCGGAAATTTCTCGACTCCCGCGTAAACCCAAACCGTCGAGGACAGGTTTTCCT 180

Qy 231 GGATGACGGTCCACAGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGCGTCCACGA 290

Db 181 GGATGACGGTCCACAGGTGTCGAGGTGTTCCATCCGGCGCATCCACCGCGTCCACGA 240

Qy 291 GGCTCATGAGCTGGTGAACGGTGGCGATCGCTACTGGGCAAGGCGCTTTTGAAGGCGAGT 350

Db 241 GGCTCATGAGCTGGTGAACGGTGGCGATCGCTACTGGGCAAGGCGCTTTTGAAGGCGAGT 300

Qy 351 TGAACACGTCAACGAAGAAATTCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCG 410

Db 301 TGAACACGTCAACGAAGAAATTCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCG 360

Qy 411 CCTCATCGAAGCAATGATCAAGCTTGATGGCACCGCCAAAGTCCCGCTGGGTGC 470

Db 361 CCTCATCGAAGCAATGATCAAGCTTGATGGCACCGCCAAAGTCCCGCTGGGTGC 420

Qy 471 AAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGCTGCTGATTCGCGAGGCGCT 530

Db 421 AAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGCTGCTGATTCGCGAGGCGCT 480

Qy 531 CCCACTGTTCGCTACATCGGTGGACCAACGACACAGCTTCTCCAGTTCCAAATGATGAA 590

Db 481 CCCACTGTTCGCTACATCGGTGGACCAACGACACAGCTTCTCCAGTTCCAAATGATGAA 540

Qy 591 CATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTGACGTTGAGGAATTCATGATCGC 650

Db 541 CATCATCAACGGTGGCGCTCACGCTGACTCCGGTGTGACGTTGAGGAATTCATGATCGC 600

651 TCCAATCGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACACGC 710

Db 601 TCCAATCGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACACGC 660

Qy 711 ACTGAAGTCCGTATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGGCTT 770

Db 661 ACTGAAGTCCGTATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGGCTT 720

Qy 771 CGCTCCTTTCGTCGGCTCCACCGCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAA 830

Db 721 CGCTCCTTTCGTCGGCTCCACCGCTGAGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAA 780

Qy 831 GGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGGCTCTGGACGTTCTTCTCTGAGTT 890

Db 781 GGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGGCTCTGGACGTTCTTCTCTGAGTT 840

Qy 891 CTTCAAGGACGGCACTTACCACTTCGAAGGTGGCGAGCACTCCGCGAGTCTGATGGCAAA 950

Db 841 CTTCAAGGACGGCACTTACCACTTCGAAGGTGGCGAGCACTCCGCGAGTCTGATGGCAAA 900

Qy 951 CGTTTACGCTGAGCTCGTTGACGCGTACCCAAATCGTCTCCATCGAGGACCCACATGACGA 1010

Db 901 CGTTTACGCTGAGCTCGTTGACGCGTACCCAAATCGTCTCCATCGAGGACCCACATGACGA 960

Qy 1011 AGATGACTGGAGGGTTTACACCACTCACCGCAACCATCGCGCAAGGTTTCAGATCGT 1070

Db 961 AGATGACTGGAGGGTTTACACCACTCACCGCAACCATCGCGCAAGGTTTCAGATCGT 1020

Qy 1071 TGGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAAGGC 1130

Db 1021 TGGGACGACTTCTTCTGTCACCAACCTTGAGCGCTGAAAGGAGGCGATCGCTAAGAAGGC 1080

Qy 1131 TGCCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCTCCACCGAGACCTTCGACGC 1190

Db 1081 TGCCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCTCCACCGAGACCTTCGACGC 1140

Qy 1191 TGTGCACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGGTGAGAC 1250

Db 1141 TGTGCACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGGTGAGAC 1200

Qy 1251 CGAGGACACCAACATTCCTGACCTCGCAGTTGCACTCAACTGTGGCGCAGATCAAGACTGG 1310

Db 1201 CGAGGACACCAACATTCCTGACCTCGCAGTTGCACTCAACTGTGGCGCAGATCAAGACTGG 1260

Qy 1311 TGCTCCAGACGTTCCGACCGCTGCGAAAGTCAACACAGCTTCTCCGATCGAGCAGCT 1370

Db 1261 TGCTCCAGACGTTCCGACCGCTGCGAAAGTCAACACAGCTTCTCCGATCGAGCAGCT 1320

Qy 1371 GCTTGGCGAGCGCGCGCTTACGCGAGGTGCGAGCGCATTTCCACGCTTTCAGGGCTAAAT 1430

Db 1321 GCTTGGCGAGCGCGCGCTTACGCGAGGTGCGAGCGCATTTCCACGCTTTCAGGGCTAAAT 1380

Qy 1431 AAAAGCGCTTTTCGACCGCGGTTAA 1455

Db 1381 AAAAGCGCTTTTCGACCGCGGTTAA 1405

RESULT 9

AX064945

LOCUS

DEFINITION

AX064945

ACCESSION

AX064945.1

VERSION

GI:12542657

KEYWORDS

SOURCE

ORGANISM

Sequence 71 from Patent WO0100844.

AX064945

Sequence 71 from Patent WO0100844.

AX064945

Sequence 71 from Patent WO0100844.

AX064945

Sequence 71 from Patent WO0100844.

AX064945

Sequence 71 from Patent WO0100844.

AX064945

Sequence 71 from Patent WO0100844.

AX064945

REFERENCE

1

Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Habberhauer, G.

Corynebacterium glutamicum genes encoding proteins involved in

carbon metabolism and energy production

Patent: WO 0100844-A 71 04-JAN-2001;

JOURNAL

REFERENCE

TITLE

AUTHORS

JOURNAL

REFERENCE

FEATURES	BASF AKTIENGESSELLSCHAFT (DE)	
	Location/Qualifiers	
Source	1..1398	
	/organism="Corynebacterium glutamicum"	
CDS	/mol_type="unassigned DNA"	
	/db_xref="taxon:1718"	
	101..1378	
	/notes="unnamed protein product; RXA00235"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAC25712.1"	
	/db_xref="GI:12542658"	
	/translation="MARIMHVFAREILDSFGNPTVRAEVLDDSGHVGAVPSGASTG VHAHEURDGGDRYLKGVILKAVENUNEEICDELAGEADDQRLLIDEMIKLDTANK SRIGANAILGVSAVAKAASAGLPLFRVYIGGNNAHLVPMNNIITGAHDSGVD VQEFMIAPIGAETFEALIRNGAEVYHAKIKEKLGSTGLDGGGFAPSVGSTRAL DLIVATEKAGFTPGKIDIALDVASSEPFKDGTYHPEGQGHAAEMANYAELVDAY PIVISEDPLQDWEVNTLATIGDKVQIVGDFFVTNPERLKEGIKAKAANSILVK VNOIGTLTETFDVADMAHRAGYTSMSHRSSETDTTIADLAVALNCGIKTGAPARS DRVAKYNQLIRIEQLLDAGVYAGRSAPFPQG"	
ORIGIN	88.5%; Score 1396.4; DB 6; Length 1398;	
	Query Match	
	Best Local Similarity 99.9%; Pred. No. 1.4e-253;	
	Matches 1397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	51	CGAAACAAGATTCTGTCGAAACAATTGGGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 110
Db	1	CGAAACAAGATTCTGTCGAAACAATTGGGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 60
Qy	111	ATAATTCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCACGT 170
Db	61	ATAATTCTAGTTAGTCCCAAGTTGGCATAGGAGCCACAGTGGCTGAATCATGCACGT 120
Qy	171	ATTCGCTCGCAAAATTTCTCGACTCCGCGGTAAACCCAAACCGTCGAGGCGAGGTTTTCT 230
Db	121	ATTCGCTCGCAAAATTTCTCGACTCCGCGGTAAACCCAAACCGTCGAGGCGAGGTTTTCT 180
Qy	231	GGATGACGGTTCCACCGGTGTCGACGGTGGCGATGCTTCCATCCGGCGCATCCACCGCGTCCAGGA 290
Db	181	GGATGACGGTTCCACCGGTGTCGACGGTGGCGATGCTTCCATCCGGCGCATCCACCGCGTCCAGGA 240
Qy	291	GGCTCATGAGTGGTGGTGGCGATGCTTACCTGCGGCAAGGCGGTTTTGAAGGCGAGT 350
Db	241	GGCTCATGAGTGGTGGTGGCGATGCTTACCTGCGGCAAGGCGGTTTTGAAGGCGAGT 300
Qy	351	TGAAACCGTCAACGAAGAAATCGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCG 410
Db	301	TGAAACCGTCAACGAAGAAATCGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCG 360
Qy	411	CCTCATCGACGAAATGATCAAGCTTGATGGGACCGCCCAACAGTCCCGCTGGGTGC 470
Db	361	CCTCATCGACGAAATGATCAAGCTTGATGGGACCGCCCAACAGTCCCGCTGGGTGC 420
Qy	471	AAACGCAATCCTTGTTGTTTCATCGCTGTTGCAAAAGGCTGCTGCTGATTCCGAGGCGCT 530
Db	421	AAACGCAATCCTTGTTGTTTCATCGCTGTTGCAAAAGGCTGCTGCTGATTCCGAGGCGCT 480
Qy	531	CCCAGTTCGCTACATCGGTGGACCAAAACGACACAGTTCCTCAGTTCCAAATGATGAA 590
Db	481	CCCAGTTCGCTACATCGGTGGACCAAAACGACACAGTTCCTCAGTTCCAAATGATGAA 540
Qy	591	CATCATCAACGGTGGGCTCAGCTGACTCCGGTGTGACGTTTCAGGTAATCATGATCGC 650
Db	541	CATCATCAACGGTGGGCTCAGCTGACTCCGGTGTGACGTTTCAGGTAATCATGATCGC 600
Qy	651	TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGGAGGTTTACCAAGC 710
Db	601	TCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGGAGGTTTACCAAGC 660
Qy	711	ACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTT 770
Db	661	ACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTT 720

Qy	771	CGCTCCTTCGTCGGTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAA 830
Db	721	CGCTCCTTCGTCGGTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAA 780
Qy	831	GGCTGGCTTCACCCAGGCAAGACATCGCTCTTGCTCTGGAAGCTTTCCTCTGAGTT 890
Db	781	GGCTGGCTTCACCCAGGCAAGACATCGCTCTTGCTCTGGAAGCTTTCCTCTGAGTT 840
Qy	891	CTTCAAGGAGGCACTTACCACTTGAAGGTGGCCAGCACTCCGAGCTGAGATGCAAA 950
Db	841	CTTCAAGGAGGCACTTACCACTTGAAGGTGGCCAGCACTCCGAGCTGAGATGCAAA 900
Qy	951	CGTTTACGCTGAGCTGTTGAGCGGTACCAATTCGTCTCATCGAGGACCCACTGCAGGA 1010
Db	901	CGTTTACGCTGAGCTGTTGAGCGGTACCAATTCGTCTCATCGAGGACCCACTGCAGGA 960
Qy	1011	AGATGACTGGAGGGTTTACACCAACCTCAACCGCAACATCGGGGCAAGGTTTCAGATCGT 1070
Db	961	AGATGACTGGAGGGTTTACACCAACCTCAACCGCAACATCGGGGCAAGGTTTCAGATCGT 1020
Qy	1071	TGGCGAGCACTTCTTCGTCAACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAAGGC 1130
Db	1021	TGGCGAGCACTTCTTCGTCAACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAAGGC 1080
Qy	1131	TGCCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCCCTCACCGAGACCTTTCGACGC 1190
Db	1081	TGCCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCCCTCACCGAGACCTTTCGACGC 1140
Qy	1191	TGTCGACATGGCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTTCCGGTGAGAC 1250
Db	1141	TGTCGACATGGCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTTCCGGTGAGAC 1200
Qy	1251	CGAGGACACCACTTGTCTGACCTCGCAGTTGCACTCAACTGTCGCGCAGATCAAGACTGG 1310
Db	1201	CGAGGACACCACTTGTCTGACCTCGCAGTTGCACTCAACTGTCGCGCAGATCAAGACTGG 1260
Qy	1311	TGCTTCAGACAGTTTCGACCGCTGTCGAAAGTAAACACAGCTTCTCCGATCGAGCAGCT 1370
Db	1261	TGCTTCAGACAGTTTCGACCGCTGTCGAAAGTAAACACAGCTTCTCCGATCGAGCAGCT 1320
Qy	1371	GCTTGGCGACCGCGGGCTTACGAGGTGCGAGGATTCGACGCTTTCCAGGCTTTTCCAGGCTTAAAT 1430
Db	1321	GCTTGGCGACCGCGGGCTTACGAGGTGCGAGGATTCGACGCTTTCCAGGCTTTTCCAGGCTTAAAT 1380
Qy	1431	AAAAGCGCTTTTCGAGCG 1448
Db	1381	AAAAGCGCTTTTCGAGCG 1398
RESULT 10		
BD163286	LOCUS	BD163286
DEFINITION	Novel polynucleotide.	
ACCESSION	BD163286	
VERSION	BD163286.1 GI:27869050	
KEYWORDS	JP 2002191370-A/1085.	
SOURCE	unidentified	
ORGANISM	unclassified	
REFERENCE	1 (bases 1 to 1275)	
AUTHORS	Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.	
TITLE	Novel polynucleotide	
JOURNAL	Patent: JP 2002191370-A 1085 09-JUL-2002;	
COMMENT	OS Corynebacterium glutamicum PN JP 2002191370-A/1085 PD 09-JUL-2002 PF 15-DEC-2000 JP 2000405096 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, KEIKO OCHIAI, PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO	

```

PI OZAKI
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00, PC
C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
PC 04,C12P13/08,
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53, PC
G01N33/566,
PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
PC C12N15/00,
PC C12N5/00,C12N15/00
CC Novel polynucleotide
FH Key Location/Qualifiers
FT source 1..1275
FT 1.1275
FEATURES
    source Location/Qualifiers
    1..1275
    /organism='Corynebacterium glutamicum'.
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"

ORIGIN
Query Match 80.8%; Score 1275; DB 6; Length 1275;
Best Local Similarity 100.0%; Pred. No. 1.2e-230;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 GTGCTGAAATCATGACGATTCGCTCGCGAAATTCGACTCCCGGGTAACCCAAACC 210
Db 1 GTGCTGAAATCATGACGATTCGCTCGCGAAATTCGACTCCCGGGTAACCCAAACC 60

Qy 211 GTCAGGACAGAGTTCCTGGATGACGGTTCCACGGTGTCCAGAGTGTTCATCCGGC 270
Db 61 GTCAGGACAGAGTTCCTGGATGACGGTTCCACGGTGTCCAGAGTGTTCATCCGGC 120

Qy 271 GCATCCACCGGTCACGAGGCTCATGAGTGTGACGGTGTGACGGTGTGACGGTGTGACGG 330
Db 121 GCATCCACCGGTCACGAGGCTCATGAGTGTGACGGTGTGACGGTGTGACGGTGTGACGG 180

Qy 331 AAGGCGCTTTTGAAGGAGTGTGAAAGCTCAAGAAAGAAATCGGCGACGAGCTCGCTGGC 390
Db 181 AAGGCGCTTTTGAAGGAGTGTGAAAGCTCAAGAAAGAAATCGGCGACGAGCTCGCTGGC 240

Qy 391 CTAGAGGCTGACGATGAGCGCTCATCGAAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db 241 CTAGAGGCTGACGATGAGCGCTCATCGAAGCAATGATCAAGCTTGTATGGCACCGCC 300

Qy 451 AACAGTCCCGCTGGTGGTGCAGCAATCTTGGTGTTCATGGCTGTGCAAGGCT 510
Db 301 AACAGTCCCGCTGGTGGTGCAGCAATCTTGGTGTTCATGGCTGTGCAAGGCT 360

Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCGGCTACATCGGTGACCAAGACGACAGCTT 570
Db 361 GCTGCTGATTCGCGAGGCTCCCACTGTTCGGCTACATCGGTGACCAAGACGACAGCTT 420

Qy 571 CTTCCAGTTCGAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCGGTGTTGAC 630
Db 421 CTTCCAGTTCGAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCGGTGTTGAC 480

Qy 631 GTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 540

Qy 691 GGCAGGAGGTCTACCAAGCACTCAAGTTCGCTCATCAAGGAAAGAGGCGCTGTCCACCGGA 750
Db 541 GGCAGGAGGTCTACCAAGCACTCAAGTTCGCTCATCAAGGAAAGAGGCGCTGTCCACCGGA 600

Qy 751 CTTGGCGATGAGGCGGCTTCGCTTCCTTCGCTGGGCTCCACCGCTGAGGCTCTTGACCTT 810
Db 601 CTTGGCGATGAGGCGGCTTCGCTTCCTTCGCTGGGCTCCACCGCTGAGGCTCTTGACCTT 660

Qy 811 ATCGTTGAGGAATTCGAGAGGCTGGCTTCACCCACGAGCAAGGACATCGCTCTTGTCTGT 870
Db 661 ATCGTTGAGGAATTCGAGAGGCTGGCTTCACCCACGAGCAAGGACATCGCTCTTGTCTGT 720

```

```

Qy 871 GACGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTTACCATTCTGAAGTGGCCAGCAC 930
Db 721 GACGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTTACCATTCTGAAGTGGCCAGCAC 780

Qy 931 TCCGACGCTGAGATGGCAACGTTTACGCTGAGCTGCTGTGACGGTACCCCAATCGTCTCC 990
Db 781 TCCGACGCTGAGATGGCAACGTTTACGCTGAGCTGCTGTGACGGTACCCCAATCGTCTCC 840

Qy 991 ATCGAGCACCACTGACGAGGAGATGACTGGGAGGTTTACCAACCACTTACCGCAACCATC 1050
Db 841 ATCGAGCACCACTGACGAGGAGATGACTGGGAGGTTTACCAACCACTTACCGCAACCATC 900

Qy 1051 GGCACAAAGTTTCAAGTCTGTCGACGACTTCTTCGTCCACCAACCTTACCGCAACCATC 1110
Db 901 GGCACAAAGTTTCAAGTCTGTCGACGACTTCTTCGTCCACCAACCTTACCGCAACCATC 960

Qy 1111 GAGGCAATCGTAAAGAGGCTGCCAACTCCATCTCTGTTTAAAGTGAACCAAGTACCGTACC 1170
Db 961 GAGGCAATCGTAAAGAGGCTGCCAACTCCATCTCTGTTTAAAGTGAACCAAGTACCGTACC 1020

Qy 1171 CTACCCGAGACCTTCCGACGCTGTCGACATGCTCACCAGGCTACACCTCCATGATG 1230
Db 1021 CTACCCGAGACCTTCCGACGCTGTCGACATGCTCACCAGGCTACACCTCCATGATG 1080

Qy 1231 TCCACCGGTTCCGCTGAGACCGAGGACACCACTTCTGACCTCGCAGTTTGCACCTCAAC 1290
Db 1081 TCCACCGGTTCCGCTGAGACCGAGGACACCACTTCTGACCTCGCAGTTTGCACCTCAAC 1140

Qy 1291 TGTGGCCAGATCAAGACTGCTGCTCCAGCAGCTTCCAGCGTGTCCGAAAGTACAACCG 1350
Db 1141 TGTGGCCAGATCAAGACTGCTGCTCCAGCAGCTTCCAGCGTGTCCGAAAGTACAACCG 1200

Qy 1351 CTTCTCCGATCCGACGAGCTGCTTGGCAGCGCGGCTTACGCGAGGTCCGACGGCATTC 1410
Db 1201 CTTCTCCGATCCGACGAGCTGCTTGGCAGCGCGGCTTACGCGAGGTCCGACGGCATTC 1260

Qy 1411 CCACGCTTTCAGGGC 1425
Db 1261 CCACGCTTTCAGGGC 1275

RESULT 11
AX1211169
LOCUS 1275 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1085 from Patent EP1108790.
ACCESSION AX1211169
VERSION AX1211169.1 GI:14037884
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
Novel polynucleotides
Patent: EP 1108790-A 1085 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
    source Location/Qualifiers
    1..1275
    /organism="Corynebacterium glutamicum"
    /mol_type="unassigned DNA"
    /db_xref="taxon:1718"

ORIGIN
Query Match 80.8%; Score 1275; DB 6; Length 1275;
Best Local Similarity 100.0%; Pred. No. 1.2e-230;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 GTGCTGAAATCATGACGATTCGCTCGCGAAATTCGACTCCCGGGTAACCCAAACC 210
Db 1 GTGCTGAAATCATGACGATTCGCTCGCGAAATTCGACTCCCGGGTAACCCAAACC 60

```


QY 211 GTGAGGACAGGCTTTCTCGGATGACGGTTCCACGGTCTCGCAGGTGTTCCATCCGC 270
Db 61 GTGAGGACAGGCTTTCTCGGATGACGGTTCCACGGTCTCGCAGGTGTTCCATCCGC 120
QY 271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACGGTGGCGATCGCTACCTGGC 330
Db 121 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACGGTGGCGATCGCTACCTGGC 180
QY 331 AAGGCGCTTTGAAGCGAGTTGAAACGTCACGTAAGAAATCGGCGACGAGCTCGTGCG 390
Db 181 AAGGCGCTTTGAAGCGAGTTGAAACGTCACGTAAGAAATCGGCGACGAGCTCGTGCG 240
QY 391 CTAGAGCTGACGATGAGCGCTCATCGAGAGCAATGATCAAGCTTGATGCGACCGC 450
Db 241 CTAGAGCTGACGATGAGCGCTCATCGAGAGCAATGATCAAGCTTGATGCGACCGC 300
QY 451 AACAAATCCGCGCTGGGTGAAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGCT 510
Db 301 AACAAATCCGCGCTGGGTGAAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGCT 360
QY 511 GTGCTGATTCGAGCGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 570
Db 361 GTGCTGATTCGAGCGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 420
QY 571 CTTCCAGTTCCATGATGAACATCATCAAGGTGGCGCTCAAGCTGACTCGGTGTTGAC 630
Db 421 CTTCCAGTTCCATGATGAACATCATCAAGGTGGCGCTCAAGCTGACTCGGTGTTGAC 480
QY 631 GTTCAGAAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCGAGGCTCTCCGCAAC 690
Db 481 GTTCAGAAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCGAGGCTCTCCGCAAC 540
QY 691 GGCAGGAGTCTACACGCACTGAAGTCCGTATCAAGGAAAGAGGCTGTCCACCGGA 750
Db 541 GGCAGGAGTCTACACGCACTGAAGTCCGTATCAAGGAAAGAGGCTGTCCACCGGA 600
QY 751 CTTGGCGATGAGGCGGCTCGCTCTTCCGTGGCTCCACCGTGAGGCTCTTGACCTT 810
Db 601 CTTGGCGATGAGGCGGCTCGCTCTTCCGTGGCTCCACCGTGAGGCTCTTGACCTT 660
QY 811 ATCGTTGAGGCAATCGAAGGCTGGCTTACCCCGAGGAGGATCGCTCTTGTCTGTG 870
Db 661 ATCGTTGAGGCAATCGAAGGCTGGCTTACCCCGAGGAGGATCGCTCTTGTCTGTG 720
QY 871 GACGTGCTTCTGATGTTCTCAAGAGACGGCACTTACACTTCGAAAGGTGGCCAGAC 930
Db 721 GACGTGCTTCTGATGTTCTCAAGAGACGGCACTTACACTTCGAAAGGTGGCCAGAC 780
QY 931 TCCGACGTGATGAGGCGGCTTACGCTGAGCTCGTTGACGGGTACCCAACTCGTCTCC 990
Db 781 TCCGACGTGATGAGGCGGCTTACGCTGAGCTCGTTGACGGGTACCCAACTCGTCTCC 840
QY 991 ATCGAGGACCCATCGCAGGAGATGACTGGGAGGTTTACACCAACCTCACCGCAACCATC 1050
Db 841 ATCGAGGACCCATCGCAGGAGATGACTGGGAGGTTTACACCAACCTCACCGCAACCATC 900
QY 1051 GGCACAAAGTTGATGTTGGCGACGACTTCTTCTGTCACCAACCTTGAGCGGCTGAAG 1110
Db 901 GGCACAAAGTTGATGTTGGCGACGACTTCTTCTGTCACCAACCTTGAGCGGCTGAAG 960
QY 1111 GAGGCGATCGCTAAGAGGCTGCCAATCCCACTCGTTAGGTTAGGTTAGGTTAGGTTAGG 1170
Db 961 GAGGCGATCGCTAAGAGGCTGCCAATCCCACTCGTTAGGTTAGGTTAGGTTAGGTTAGG 1020
QY 1171 CTCACCGAGACCTTCGAGCTGTGCACATGCTCACCGCGAGGCTACACCTCCATGATG 1230
Db 1021 CTCACCGAGACCTTCGAGCTGTGCACATGCTCACCGCGAGGCTACACCTCCATGATG 1080
QY 1231 TCCACCGCTTCGGTGAGACCGAGGACACCAACCATTTGCTGACCTCGAGTTGCACTCAAC 1290
Db 1081 TCCACCGCTTCGGTGAGACCGAGGACACCAACCATTTGCTGACCTCGAGTTGCACTCAAC 1140

QY 1291 TGTGGCCAGATCAAGACTGGTCTCGACGACGTTCCGACCGTCTCGAAAGTACAACCCAG 1350
Db 1141 TGTGGCCAGATCAAGACTGGTCTCGACGACGTTCCGACCGTCTCGAAAGTACAACCCAG 1200
QY 1351 CTTTCTCCGATCGAGCAGCTGTTGGCGACGCGGCTTACGACGTTGCGACGGTTCGACGGCATTC 1410
Db 1201 CTTTCTCCGATCGAGCAGCTGTTGGCGACGCGGCTTACGACGTTGCGACGGTTCGACGGCATTC 1260
QY 1411 CCACGCTTTCAGGCG 1425
Db 1261 CCACGCTTTCAGGCG 1275

RESULT 12
AP005217
LOCUS 300750 bp DNA linear BCT 24-JUL-2003
DEFINITION Corynebacterium efficiens YS-314 DNA, complete genome, section 4/11.
ACCESSION AP005217 BA000035
VERSION AP005217.1 GI:23492722
KEYWORDS Corynebacterium efficiens YS-314
SOURCE Corynebacterium efficiens YS-314
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
AUTHORS Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K. and Gojobori, T.
TITLE Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of Corynebacterium efficiens
JOURNAL Genome Res. 13 (7), 1572-1579 (2003)
MEDLINE 22723752
PUBMED 12840036

REFERENCE 2 (bases 1 to 300750)
AUTHORS Kawarabayasi, Y., Yamazaki, J., Hino, Y., Kikuchi, H. and Director-General of Biotechnology Center.
TITLE Director-Submission
JOURNAL Submitted (17-MAY-2002) Director-General of Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; Nishihara 2-49-10, Shibuya-ku, Tokyo 151-0066, Japan (E-mail:bioente.go.jp, Tel:81-3-3481-1933, Fax:81-3-3481-8424) Kawarabayasi, Y. is officially affiliated with the National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, 305-8566 Japan
Nakamura, Y., Ikeo, K., Suzuki, M. and Mashima, J. are at the National Institute of Genetics, Mishima, Shizuoka, 411-8540 Japan
Itoh, T. is at the Japan Biological Information Research Center, Koto-ku, Tokyo, 135-0064 Japan
Yamagishi, A. is at Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo, 192-0392 Japan
Nishio, Y., Usuda, Y. and Sugimoto, S. are at the Ajinomoto Co., Inc., Kawasaki, Kanagawa, 210-8681 Japan
The other authors are at the National Institute of Technology and Evaluation, Shibuya-ku, Tokyo, 151-0066 Japan.

FEATURES
source
1. 300750
/organism="Corynebacterium efficiens YS-314"
/mol_type="genomic DNA"
/strain="YS-314"
/db_xref="taxon:196164"
252..620
/note="CE0885"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC17695.1"
/db_xref="GI:23492723"
/translation="MDQAPSGCLVDCVHGVIYDTPVPLLFPVGTSGNKRKIVLDLNF
HYWLNKPKRLLTFLWSGCGCLFGWIGLFTPWHTVYVVCVFCGDFDPSGRLLAAC
LTHASRTMKPCLQGLVANG"
529..1936

CDS
529..1936

gene

```

/ gene="rRNA_16S"
529..11936
/ gene="rRNA_16S"
/ product="16S ribosomal RNA"
/ note="CE:RNA01"
complement(2042..2215)
/ note="CE0886"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17696.1"
/ db_xref="GI:23492724"
/ translation="MSSQTNSMPPTVHHHLGDVFADPWWMIHPITTCWQPHPGSNT
NNPHEHSPWVY"
2399..5481
/ gene="rRNA_23S"
2399..5481
/ gene="rRNA_23S"
/ product="23S ribosomal RNA"
/ note="CE:RNA02"
5606..5725
/ gene="rRNA_5S"
5606..5725
/ gene="rRNA_5S"
/ product="5S ribosomal RNA"
/ note="CE:RNA03"
5925..7088
/ note="CE0889, similar to AX063829-1|CAC25156.1| percent
identity: 79 in 387 aa"
/ codon_start=1
/ transl_table=11
/ product="putative aminotransferase"
/ protein_id="BAC17699.1"
/ db_xref="GI:23492727"
/ translation="MRNKYVVERLRPGETIFATMTORANEAGAINLQGGFPDGDGA
PMLEASQQLGGNNQYSGGRGDIGLRTAVASDRHRKFLDYDPTDLITVGATEAI
SASVLGVPQGVLEVPYDYDAAYAAAGATRAVAPLKEEDNTWVLDTDVHAHAV
AKPTSMIVNSPHSPGTSVFSKRLEALAGIARAYDILLVLSDEYHLLTDFDKVSHTA
V
TSLGMDWRTITVSSAAKTENVTKGTMALAPALEAVIRAKQFMSYVGATPQPA
VAHAHEKPKWKMRGLONKRDILTDALTQAGLVKVDHSGTTFVADIGERGAEF
CFDLDRGVAAIPQAVFDKPEQWSSKRFAPFCKQEDTLEAAQRLRAAGSL"
complement(7180..7821)
/ note="CE0890"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17700.1"
/ db_xref="GI:23492728"
/ translation="MTHFSFDVDEAYARKNNEFLRDKRLQISAFFGLVLLAIGGVL
YVLADGAVGMWMLIVMGIMAAISFVMI PVLPROVGSASLNTYPLVPVAVAEINPR
DVLLALVNTNVDTLPPRWGLAVRTINRLGVHDKLGERVPSVAVSGRTIRDDHW
DEISPMITWGTDDVVRRAEKTIPHDLWNKLEKNRRLDEVKNTMNLVL"
complement(7822..9495)
/ note="CE0891, similar to AL583924-93|CAC31112.1| percent
identity: 75 in 547 aa"
/ codon_start=1
/ transl_table=11
/ product="putative DNA helicase"
/ protein_id="BAC17701.1"
/ db_xref="GI:23492729"
/ translation="MAPGCDGLIVQSDKTVLLETCHDLAGEARAALAPFAELERAPHE
VHTYRITPLAMNARAAGHDAQVQVLEYSRFPVQPLLDVDAETMSRYGRVRMHK
HPAGLILESSEPAILVEISPHRKIKMPLGOOLDPETIAVHPSERGLKQELLKIGWP
AEDLAGYVDEAHFIALSTEVEDSLRDYQQAADSFWEGSGVVPLPCGAKGTWGA
ASMAQAQTLLITNTVAGWQKDELLRTLTLEDEIGSEYGERKEIRPVITATYQV
VTRTKGEYKALELFDPRWGLIYDEVHLLPAPVFRMTSLQSRRLGLTATLVRED
RGDGVFLSPKRPDAPFWKOLEAQGFATACVEVTRTMDAERMVYATAETADRYR
LAATAHTKTRVVRILECHQOQPTLIIGAYLDLQEEELGAEFDAPIVDGKTSNKKRGL
FDRFRSGLNVLVVSKVANFSIDLPEAAVAIQVSGTFGSQEEAQRGLLRPRKADGG
EAHPYSVSRDRLDTEVAHRQRFLAEGQGVAYRIIDADDLLFPITEKET"
complement(9548..11851)
/ note="CE0892, similar to AE006976-11|AAK45126.1| percent
identity: 36 in 741 aa"

```

```

/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="BAC17702.1"
/ db_xref="GI:23492730"
/ translation="MPCFPFSQIRQTKVTSTAPKLGIMKDKDSPVTLTGWLEQDDDD
QLSGILNRDPTALPLPPGLSGLAARLQLRASIQRAVHKLSALELAVLEAAANAGAE
NPITAPVEVRLHRLHSAAPTAEQIGALNRLQLALVFGGERFIILPETWATLP
TNQQLPEEGRSLTFGEASSTALPARHKILVTLVTSNGLGLTRDAALSADPARP
IPOLIKAGLLSRVDEQTVKLPGVVVRVLEGGEDTAMRVVPVPSAPGADAGIAAGL
EVARLQMLDALSTAPASTLKSGALGVWRVTLTRELQVDEQOQVAVSVLCVSSGLL
RRGVPDLPADDGGDYIAPTVADEMLQATLAEQLAHLMRGWTQIYAALVGEADE
KNPIHLLSPASRRDALPETRAMILQSLTRVGEEDLEADLFHFRPLAASRTPRETIH
LLBEARWIGAVAGGVTSPTARLTQAPAGVPEIADIADVAVRPKPYDYFTQADYTV
MVPGLQPEMQIKIGQLADLESGLASVYRISEASLRHAMDGLTATEIEDFLQHSVA
TGLPQSVAYILLQDIARHGTLRGSPALCYLRSDDPALLHSAVEAAGLGRQIAPTVA
VSNASLIEVITALKRAGMQPVAEDALGASLDLAPRPVAPAAEQPRASGPDDESIRA
AVAAIRREEVARTGTSDTLAVLQAAVRGQRTVLIGFVDKQGVAVHRTVKPLTVNA
GQVDALDESTGAVHRFMLHRITEVID"
11836..12036
/ note="CE0893, similar to AL583924-91|CAC31110.1| percent
identity: 50 in 62 aa"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="BAC17703.1"
/ db_xref="GI:23492731"
/ translation="MGNMANVEKKGYVDPAWPDHPDSEGHVVVTELIAPYACASSPWGD
DMEPPVPAETIGYVHPYTRNR"
complement(12229..12822)
/ note="CE0894, similar to AX133781-1|CAC38957.1| percent
identity: 71 in 195 aa"
/ codon_start=1
/ transl_table=11
/ product="conserved hypothetical protein"
/ protein_id="BAC17704.1"
/ db_xref="GI:23492732"
/ translation="MLFMRHSKATTPFVKMAATTVAFGAAVAFAPAAASAAPDSW
DRLANCSGNGWAINTCNGYHGGLQFSPSTWRAYGDFEFPAYQATREQQIIAUAERT
LAGQMGWAPACSKLGLNSAPTRNLNPAPAPAPAPAPAEYRATVDTNTNPVVGSS
DLNTVQGIYDAVSGTLAQYGVQVPAEIQAHYNAFGR"
13342..13725
/ note="CE0895, similar to AE006977-4|AAK45135.1| percent
identity: 52 in 127 aa"
/ codon_start=1
/ transl_table=11
/ product="putative cold shock protein"
/ protein_id="BAC17705.1"
/ db_xref="GI:23492733"
/ translation="MPVGTWKYDPERGFGFVSNPBGDCDFVGKQVLPKGVTELHQGQ
RIEFDPAAGRKGQALRVKYLETPRRRPQHTYKPEELNGIISDMVTLMLEGTVPALLK
GHPDRKVGQVAKILRAVAKELEA"
complement(13881..14408)
/ note="CE0896"
/ codon_start=1
/ transl_table=11
/ product="hypothetical protein"
/ protein_id="BAC17706.1"
/ db_xref="GI:23492734"
/ translation="MASRKTQRNFIQIATLLVAVVVVVISVVVFQNMNNRPEPEPE
TVSITASSPAGEIEVFPYSICEPGVCAENEVPTLEVGADELRLLTPEIITHDHWYL
LTYVDDPAANDFEVFTSYETTEVTVPGSVDVPTGGERPRLVVVEVTSVMIGTDNGE
ETPVTTWTSATVEQ"
complement(14424..15254)
/ note="CE0897, similar to AE001874-1|AAAF09704.1| percent
identity: 40 in 267 aa"
/ codon_start=1
/ transl_table=11
/ product="putative glutamine cyclotransferase"
complement(15478..16051)
/ note="CE0898, similar to AE001874-1|AAAF09704.1| percent
identity: 40 in 267 aa"

```

Query Match 65.7%; Score 1036.4; DB 1; Length 300750;
 Best Local Similarity 81.1%; Pred. No. 9.8e-186;
 Matches 1235; Conservative 0; Mismatches 246; Indels 41; Gaps 1;

QY	96	CATTGATCAGTGAATAATCTAGTTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGC	155
Db	170262	CCITTTGTGCGGTGAACATATCTGGTTAGTTCCCAAGTTGGCATAGGAGGCCACAGTGGC	170321
QY	156	TGAAATCATGCATGATATTCGCTCGCGAAATTTCTCGACTCCGCGGTAAACCCAAACCGTCTGA	215
Db	170322	TGAAATCATGCATGATATTTGCTCGCGAGATCATGGATTCGCGCGGTAAACCCGACCGTCTGA	170381
QY	216	GGCAGAGGTTTCTGATGACGTTTCCACGCTGTCGCAAGTGTTCATCCGCGCATC	275
Db	170382	GGCAGAGGTTTCTGATGACGTTTCCACGCGGTTCGCGCGTTCCTCCGAGCGTC	170441
QY	276	CACCGCGTCCAGAGGCTCATGAGCTGCGTGAGGTTGGGATCGCTACCTGGGCAAGG	335
Db	170442	CACCGGTGTCCAGAGGCGCACGAGCTGCGTGAGCGGCGAGATCGTTACCTGGGCAAGG	170501
QY	336	CGTTTGAAGCGAGTTGAAACAGTCAACGAAGAAATCGGCGACGAGCTCGTGGCCTAGA	395
Db	170502	TGTTCTCAAGGCTGTTGAGAACGTTCAACGAGGAGATCGGCGACGAGCTGGCGGCTTGA	170561
QY	396	GGCTGACGATCAGCGCTCATGCAAGAACATGATCAAGCTTGATGGCACCGCAACAA	455
Db	170562	GGCGACGATCAGCGTCTCATCGATGCGGCGCATGATCAAGCTGGAACGCAACCGAAGAACAA	170621
QY	456	GTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGAAAGGCTGCTGC	515
Db	170622	GTCCCGCTGGGCGCAACGCAATCTTGGTGTTCATGGCTGTGGCGCAAGGCGCGC	170681
QY	516	TGATTCGCGAGGCTCCCACTGTTTCGCTACATCGGTGGAACCAACGACACGCTTCTTCC	575
Db	170682	GGATTCGCGAGGCTCGCCCTGTTCCGTTACATCGGTGGCGCGCAACGCGCAAGTGTGCTGC	170741
QY	576	AGTTCCAATGATGAACATCATCAACGCTGGCGCTCAACGCTGATCTCCGGTGTGACGTTCA	635
Db	170742	CGTCCGATGATGAACATCATCAACGCTGGCGCACACGCGCATCTCCGGTGTGATGTCCA	170801
QY	636	GGAAATCATGATCGCTCCAATCGGTGCAGACCTTCTTGAGGCTCTCCGCAACGCGC	695
Db	170802	GGAGTTTATGATGCCCGCATCGGTTCGATCTCTTCGAGGCGCTTCGCGGAGCGC	170861
QY	696	GGAGGTCTACACGCACTGAAGTCCGTCAACGAAGAAAGGCGCTGCCACCGGACTTGG	755
Db	170862	GGAGGTCTACACGCACTGAAGAGGTCAACGAAGAGGCGCTGCCACCGGTCTGG	170921
QY	756	CGATGAGGCGGCTTCGCTCTTCGCTCGGCTCAACCGGTGAGGCTCTTGAACCTTATCGT	815
Db	170922	CGATGAGGCGGCTTCGCTCTTCGCTCGGCTCAACCGGTGAGGCTCTTGAACCTTATCGT	170981
QY	816	TGAGGCAATCGAAGGCTCGCTTCAACCGGCAAGGACATCGCTCTTCTGGAAGT	875
Db	170982	CGAGGCAATCAAGAGCGCGCTTCGAGCGGGCAAGGACATCGCTCTGCGCTGATGT	171041
QY	876	TGCTTCTCTGAGTTCTTCAAGAGCGCACCTACCACTTCGAAGGTGCCAGCACTCCGC	935
Db	171042	TGCTTCTCTGAGTTCTTCAAGAGCGCAATACCACTTCGAGGTTGCCAGCACACGC	171101
QY	936	AGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGAGCGGTACCCAACTGCTTCCATCGA	995
Db	171102	CGAGGAGATGGCAACGCTTACGCGCGAGCTTGTGACGAGTACCCGATCGTCTCCATCGA	171161
QY	996	GGACCCACTCGAGGAAGATGACTGGGAGGGTTACACCAACTCAACGCAACCACTCGGCGA	1055
Db	171162	GGATCCGCTCGAGGAGACGATGGGATGGTACGTCGCTGACCTGACCGCCGATCGGTGA	171221
QY	1056	CAAGGTTTCAGATCGTTGGCGACGATCTTCTTCGTCACCAACCTGAGCGCTCGAAGGAGG	1115
Db	171222	CAAGGTTTCAGATCGTTGGCGATGACTTCTTCGTCACCAACCTGAGCGCTCGAAGGAGG	171281
QY	1116	CATCGCTTAAGAGGCTGCCAACTCCATCTGGTTAGGTTGAACGAGATCGGTACCTTCAC	1175
Db	171282	CATCGCCCAAGAGCGCCCAACTCCATCTGGTCAAGGTCAACGAGATCGGTACCTTCAC	171341

QY	1176	CGAGACCTTCGACGCTGTGCAATGCTACCGCGCAGGCTACACCTCCATGATGTCCCA	1235
Db	171342	CGAGACCTTCGACGCGCTGTGCAATGCGCACCGCGCAGGTTACACCTCCATGATGTCCCA	171401
QY	1236	CCGTTCCGGTGAGACCGGAGGACACCACTTGTCTGACCTCGCAGTTTGCACCTCAACTGTGG	1295
Db	171402	CCGTTCCGGTGAGACCGGAGGACACCACTTGTCTGACCTCGCAGTTTGCACCTCAACTGTGG	171461
QY	1296	CCAGATCAAGACTGGTGTCTCCAGCAAGTTCGCAACGCTGTCCGCAAGTACCAACAGCTTCT	1355
Db	171462	CCAGATCAAGACCGGTGACACCGCACGTTCCGACCGTGTGGCCAAAGTACCAACAGCTGCT	171521
QY	1356	CCGATCGACAGCTGCTTGGCGACCGCGGCTTACGCGAGTTCGAGCGCATTCGCCAG	1415
Db	171522	CGCATCGACAGCTGCTTGGGTGACCGACCGCTTACGCGCGGCTTCCCGGTTCCCGACG	171581
QY	1416	CTTTCAGGGGTAAAA-----TAAAA	1434
Db	171582	TTTCCAGGGCTAGACGCTTTTCAGCAACCCCTGACCCCGCATCCCGTCACTTCCCTCAA	171641
QY	1435	GCCTTTTCGACCGCGGTAACTCAAGGTTGCGGCGCTCGTTGCTTACTACTGTTC	1494
Db	171642	AGTGGCGGATGTTGGCTATCGGGCATGTTTCGGGCGTTCGCAAGCAGTGGCGTGAC	171701
QY	1495	TGCTGTGACTATCATGAGGATTCGCAAGCAGAGAAAATCTATAAAGGCTTGTTC	1554
Db	171702	GGTGGGACTATGATGTTGATTTATGGCAAGCAGAGAAAATCTATAAAGGCTTGTTC	171761
QY	1555	CTGCTCAAGCAGGGAACGTGC	1576
Db	171762	CGCTCTCAGCAGAGATCGTGC	171783
RESULT 13			
BX248356			
LOCUS	347625 bp	DNA	linear BCT 06-NOV-2003
DEFINITION	Corynebacterium diphtheriae gravis NCTC13129, complete genome;		
ACCESSION	BX248356		
VERSION	BX248356.1	GI:38199593	
KEYWORDS	complete genome.		
SOURCE	Corynebacterium diphtheriae		
ORGANISM	Corynebacterium diphtheriae		
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
AUTHORS	Corynebacterineae; Corynebacteriaceae; Corynebacterium.		
	1 (bases 1 to 347625)		
	Cerdeno-Tarraga, A.M., Efstathiou, A., Dover, L.G., Holden, M.T.G.,		
	Pallen, M., Bentley, S.D., Besra, G.S., Churcher, C., James, K.D., De		
	Zoyza, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T.,		
	Hamlin, N., Holroyd, S., Jørgensen, K., Moule, S., Quail, M.A.,		
	Rabinowitz, E., Rutherford, K., Thomson, N.R., Unwin, L.,		
	Whithead, S., and Barrell, B.G. Parkhill, J.		
	The complete genome sequence and analysis of Corynebacterium		
	diphtheriae NCTC13129		
TITLE	Nucleic Acids Res. 31 (22), 6516-6523 (2003)		
JOURNAL	14602910		
PUBMED	2 (bases 1 to 347625)		
REFERENCE	Cerdeno-Tarraga, A.M.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-OCT-2003) Cerdeno-Tarraga A.M., submitted on behalf		
JOURNAL	of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust		
	Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:		
	amct@sanger.ac.uk		
FEATURES	Location/Qualifiers		
source	1. .347625		
	/organism="Corynebacterium diphtheriae"		
	/mol_type="genomic DNA"		
	/strain="NCTC13129"		
	/db_xref="taxon:1717"		
	/note="biotype gravis"		
	complement (81. .206)		
	/locus_tag="DIP0718"		
	complement (81. .206)		
gene			
CDS			

```

/locus_tag="DIP0718"
/notes="Doubtful CDS. No significant database matches"
/codon_start=1
/transl_table=11
/product="Putative membrane protein"
/protein_id="CAE49240.1"
/db_xref="GI:38199584"
/translations="MKYRSNIIIVCSLIPQLPKARHIPAKIPSSPPVRIELP"
complement(141..206)
/locus_tag="DIP0718"
/notes="Signal peptide predicted for DIP0718 by SignalP 2.0
HMM (Signal peptide probability 0.699) with cleavage site
probability 0.667 between residues 22 and 23;
signal-peptide site"
257..484
/locus_tag="DIP0719"
257..484
/locus_tag="DIP0719"
/notes="Similar to Mycobacterium tuberculosis CDC1551
hypothetical 10.1 kDa protein MT3304 TR:AAK47646
(EMBL:AE007142) (97 aa) fasta scores: E(): 4.5e-11; 51.38%
id in 72 aa"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49241.1"
/db_xref="GI:38199585"
/translations="MDIKIGFSDSPRELVSREDHNEIVDRIHNALRDAEGVLDLVD
DHGNRYLVNRARTAYEVGTAPTVGFAGA"
487..1365
/locus_tag="DIP0720"
487..1365
/locus_tag="DIP0720"
/notes="Similar to Mycobacterium tuberculosis hypothetical
31.0 kDa protein Rv3207c or MTCY07D11.19 TR:O05859.
(EMBL:Z95120) (285 aa) fasta scores: E(): 1.7e-44; 46.69%
id in 287 aa"
/codon_start=1
/transl_table=11
/product="Putative membrane protein"
/protein_id="CAE49242.1"
/db_xref="GI:38199586"
/translations="MTDPAESFFVRFARDYGMRAYAIPVLAVITVWVLIDVFRTPAE
TTTTATVGGAPTATATSAQKGPDPARQNPDAITELPSGPFQKQEGTYRTVNA
GAGAGKHDPKVFYTVIVENGINTAAVGGDDAFAAMVDATLTNPKSWTHDKRGFEHV
DAGAKDPDLRIQLSSVDTHGLGNNIAMETSCFYIGNRVINERWYRGAKPFG
DLGAYRQYLINHEVGHGIGFANHEPCNGELAFIPMMQQTLSLSNSELFAIDANETYN
DDGAVCSANPWPYPFA"
535..603
/locus_tag="DIP0720"
/notes="1 probable transmembrane helix predicted for
DIP0720 by TMHMM2.0"
1165..1194
/locus_tag="DIP0720"
/notes="ScanRegExp hit to PS00142, Neutral zinc
metallopeptidases, zinc-binding region signature."
1389..2246
/locus_tag="DIP0721"
1389..2246
/locus_tag="DIP0721"
/notes="Similar to Mycobacterium leprae hypothetical
protein ML0818 TR:Q9CCG7 (EMBL:AL583919) (297 aa) fasta
scores: E(): 1.2e-34, 39.47% id in 266 aa"
/codon_start=1
/transl_table=11
/product="Conserved hypothetical protein"
/protein_id="CAE49243.1"
/db_xref="GI:38199587"
/translations="MELQQLQHVDRDAFAPNTPPRQLGVADYGRVGNVAFSOVTH
PDRSAWSARVDRKLPQGLRVPRPKTDGRYINAGWRANSFEGTLARVDETVAA
LRUDTALADLDVPSFSEPSHDLFSLADHSWAPEPMVALGVTSDEAVFLNPAQLT
AQRLKVLKRLRIDIQITQVTHDMFPATTIYSGNQMPVVDLVGVARPYGYTAAVCI
VDALLAGAVDEGIIDRESHTQHRDQLLLRALVYRIAVALHHPSTSTNTGNLEWVSQT
IMSRASVTL"
2250..5432
/locus_tag="DIP0722"
2250..5432
/locus_tag="DIP0722"
/notes="Similar to Streptomyces coelicolor putative
ATP-dependent DNA helicase 2SC3B6.07 TR:Q9FCK5
(EMBL:AL390968) (1159 aa) fasta scores: E(): 6.2e-29,
32.21% id in 1161 aa"
/codon_start=1
/transl_table=11
/product="Putative ATP-dependent DNA helicase"
/protein_id="CAE49244.1"
/db_xref="GI:38199588"
/translations="MDAFSTPQDSPTTSAQVIGEDLTAGIPPLQQSFVYAQRPRVEIV
YNDHSEPIREWDNQLFOERGTRWTGAAGAGVSTLLMDTVABRIRGVSPESIIVVA
ASRESAARLPAAGIAHRIADGSYTPASLVSRSLSLAFALLRSLSDEQLRLTGAQDA
VIRELQGHVDNFSLAAMPARQAQLGLIGFAKEVRDFLRSRSGLEPDLBELGA
RYGRPMWASAGKFMREYQQVMNLGGTHSLNASLVSKLFCFIDPMGMRTIIIDDAQH
LDPSQAQLQALMYTDTFTIADGPQQSFHFGASPEFLHAPVDHETLTIYSFRQP
TVBAKVLIGSTGDSQSEFVADLLRRSHLLGIRWMDMAVIVRSTADIPSLRRALLSAGVP
VQEDPSDIIISEQRIVSALILAVRAIYQELSAQELRELAGLPGIGADTVTLRLPRGL
RKAMRAGGNRAIEMRELIDPKSESEQTQREQVEAVLTDRELAVVDKIRAVLVRG
AQPSVEEILWEIWDASLSSHLLQTVSLRGVGAQADRLDADVMALFDDAAGDWERR
PTASITFSVRHIAEQELPTGVRDRRLATADAVRIVTAHSGLQQQMHTVI VAGVQEGTW
PSIQETGTLFGQBDLIDLIDSGIEPTNTYISRSAEKLEKEKRLFHVARTGRHVVITA
VSPESDTAAEPSRFQIGIAESKVQGPPTTASVQSEIDTFVHLDDQVYRLLSVPS
IVAEARELANPDSQRRRQAARLARLALHQQVGPARGPQMGYGGPSETTSDISK
VPSLEIENALLCPRLAKRLVEENPTIHMILKGTLAHAEAVGVGDVDPNEASQLVT
QAFEALLDPAWLSLPHHMSQWMLQRLAHWDVSHAQRELVGVGVNVVVAAPGVEL
RGRIDRLERNDAEFHIVDFKTVQAVTKDEANENKQLLAYQLALHGRKLMQRNGEPA
INTTAEQPLTVDAQVLVYPATDKTIVTTREQAPKDAEELEKFSATLIPALLESLSGPQ
LVARINTRCDCKIKTMCPAQPEGKMWPEC"
2250..2558
/locus_tag="DIP0722"
/notes="ScanRegExp hit to PS00430, TonB-dependent receptor
proteins signature 1."
3045..3200
/locus_tag="DIP0722"
/notes="HMMPFam hit to PF00580, UvrD/REP helicase"
3285..3431
/locus_tag="DIP0722"
/notes="Similar to Mycobacterium tuberculosis CDC1551
helicase, UvrD/Rep family MT3295 TR:AAK47638
(EMBL:AE007142) (1101 aa) fasta scores: E(): 2.9e-25,
43.03% id in 1120 aa"
/codon_start=1
/transl_table=11
/product="Putative helicase"
/protein_id="CAE49245.1"
/db_xref="GI:38199589"
/translations="MISPSQELSCALGQKFPPTQQADVSISSPLATLVVAGAGAKTE
TMAARVVLVAGLVDPDRVLGLTFTTKAAQLQSKRIRDRLEQLAGIDNLRDLPTGA
LATKLEAIAPTVSTYDSYAGRLISEYGLLLPVPSSRLISQTELFQIAHSIVSAHTGA
SPATKKELTVSFENPPELDLNRVSRRELGVDPDRPRVPQPPGPAALNGIVRLGF
LNTSNPNTVTSTLISLVSEMDNMVSPTDIEBSAAPFLAMIEDVEATSRAPSKVEY
KWHDQVLRNELLPLVQQLKTHLADNHLMTFGSQMSLAARLAAENPQVGASQRRRYQI
IMLDEYQDTGHAQRLKLLSKFAGTAVTAGDPMQSIYGRGATAANLERFLTDFGSG
FPMDBEERSVVAADLAAQVBOHGDHPFTAAVLVRKXKHAATALELQGVGPVVEIVG
LAGLIGTPEVADLVAJLTLVRYDYOAMRILAGPSVGLGMADLMALSDRAYNLSGR
DRRATTELSDRPLERLAKIITADTPSDQSDISVGLAEVADLERDSDSGDRYSAGKS
ERUTLAARURYLRITNSLSPLDFADIERVFGIRTEVIREQREDPRSDGTAHLDR
FAEVQDFSRIPGANISLLDLYSLAESENGLEPGEVQVTAARVQLTLTVHKAAGLEW
QHVAVLHADANTYVAKASTWLTNASAVPSALRGDAKGDDELVGAPVFEIDTPTDAEAL

```


FEATURES	source	Location/Qualifiers	gene	misc_feature	repeat_region	gene	misc_feature	repeat_region	
stem_loop		1. .276800 /organism="Streptomyces coelicolor A3(2)" /mol_type="genomic DNA" /strain="A3(2)" /db_xref="taxon:100226" 48. .98 /note="Score 59: 21/22 (95% matches, 0 gaps) complement(146. .652) /gene="SCO3054" /note="synonym: SCBAC19G2.09c" complement(146. .652) /gene="SCO3054" /note="SCBAC19G2.09c, conserved hypothetical protein, len: 168aa; similar to many in Streptomyces coelicolor eg. TR:Q9F2X0 (EMBL:AL392148) SCD20.09 hypothetical protein (190 aa) fasta scores; Opt: 250, Z-score: 285.6, 34.503% identity (38.816% ungapped) in 171 aa overlap." /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44519.1" /db_xref="GI:14970940" /db_xref="UniProt/TrEMBL:Q93J49" /translation="MSERSAPFGSLGVETLVNLTLETGADSLDTEGRAPPLGTLDPAAERLESRLATLLAHGHPHRAVTPIGELAAPLVAVDPAAGSAALAPAGE GSITARVAASAQAALVAGTWSRLKACEAACHWAYDPSAGRGRWCSMQVCGARAKM RYRAKES" 764. .1123 /gene="SCO3055" /note="synonym: SCBAC19G2.10" 764. .1123 /gene="SCO3055" /note="SCBAC19G2.10, conserved hypothetical protein, len: 119aa; similar to many in Streptomyces coelicolor e.g. TR:O85701 (EMBL:AF072709) 3SCF60.11c conserved hypothetical protein (131 aa) fasta scores; Opt: 322, Z-score: 387.9, 42.017% identity in 119 aa overlap and downstream neighbouring CDS SCBAC19G2.11" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44520.1" /db_xref="GI:14970941" /db_xref="UniProt/TrEMBL:Q93J48" /translation="MPVAELGPVVLDCDPRLARLFYADVLGCTVGEQDWDLRPLPG GRALFAQAPGFVPKPKPAPDQSQQFHLDDLVDKDLDAEAXVLELGARPLDAEDRTGT FRVYADPAGHPFCLCAC" 786. .1114 /gene="SCO3055" /note="Directly repeated at 11832. .12181" 1153. .1545 /gene="SCO3056" /note="synonym: SCBAC19G2.11" 1153. .1545 /gene="SCO3056" /note="SCBAC19G2.11, conserved hypothetical protein, len: aa; similar to many in Streptomyces coelicolor eg. TR:Q9F3E9 (EMBL:AL450289) SC2H2.06 hypothetical protein (150 aa) fasta scores; Opt: 340, Z-score: 415.4, 46.032% identity in 136 aa overlap and upstream neighbouring CDS SCBAC19G2.10" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC44521.1" /db_xref="GI:14970942" /db_xref="UniProt/TrEMBL:Q93J47" /translation="MAPAARFRSVVDCPPRELARFYAAVGGTTPDEADPDWVVLQV PGGPRLSFORAPDLTPPEWPSDRNAQQFHLDPDGGATWAEMDAANDRVIALGARPLD LEDREDKDFWYADPAGHPFCLCRIEHT" 1178. .1527 /gene="SCO3056"							
gene		/note="Directly repeated at 11440. .11768" complement(1558. .2688) /gene="SCO3057" /note="synonym: SCBAC19G2.12c" complement(1558. .2688) /gene="SCO3057" /note="SCBAC19G2.12c, possible dipeptidase, len: 376 aa; similar to many eukaryotic peptidases e.g. SW:P31430 (MDPL_RAT) rat renal dipeptidase (410 aa) fasta scores; Opt: 440, Z-score: 474.0, 36.056% identity (38.906% ungapped) in 355 aa overlap. Also weakly similar to TR:Q9KH70 (EMBL:AF268476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; Opt: 273, Z-score: 298.4, 28.159% identity (30.116% ungapped) in 277 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase. Also similar to neighbouring upstream CDS SCBAC19G2.13c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44522.1" /db_xref="GI:14970943" /db_xref="GOA:Q93J46" /db_xref="UniProt/TrEMBL:Q93J46" /translation="MADLQDDLHTGAGGLDLPVLAEEFPPEPYAFVSPDDDEP LARAHVLAAPVADGCGNLPWALRHLPLWYDLRGLRISAVDTPVRLREGHVALLWSL LHPESLDGDRAGVATLBOLDLKVTVRAHPEGLRLAYDAQQAIDARNCGRIAVLPQPA GAALGRDLGILRSLHALGRLVLTLSQVSWASAGLTRFGEVVRNMRNLGVVADLSG ASAEVTRTFVASKAPALCTRSAAALRPHNPDLDELVELGAAGLCMVPLTASQT GPTVRVDADHLDHVRTVAGSPQVSLGSDYDGAAPHELGLDPSYPRLVLAELLRRGQD EADVALLTGNVQVRLFAAAFTAKAAQLRREPSTATIADLDG" complement(1561. .2616) /gene="SCO3057" /note="Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase, score 121.70, E-value 1.4e-32" complement(1562. .2643) /note="Degenerately, directly repeated at 13490. .14610" complement(2808. .4010) /gene="SCO3058" /note="synonym: SCBAC19G2.13c" complement(2808. .4010) /gene="SCO3058" /note="SCBAC19G2.13c, possible dipeptidase, len: 400 aa; similar to many eukaryotic peptidases e.g. SW:P31429 (MDPL_RABIT) rabbit renal dipeptidase (410 aa) fasta scores; Opt: 788, Z-score: 879.1, 42.432% identity (48.58% ungapped) in 403 aa overlap. Also weakly similar to TR:Q9KH70 (EMBL:AF268476) thermostable dipeptidase from Brevibacillus borstelensis (307 aa) fasta scores; Opt: 450, Z-score: 506.1, 31.034% identity (36.486% ungapped) in 348 aa overlap. Contains Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase. Also similar to neighbouring downstream CDS SCBAC19G2.12c" /codon_start=1 /transl_table=11 /product="putative dipeptidase" /protein_id="CAC44523.1" /db_xref="GI:14970944" /db_xref="GOA:Q93J45" /db_xref="UniProt/TrEMBL:Q93J45" /translation="MTSLEKARELLRFFVVDGHNLDLPWALREQVRYDLDARDIAADQ SAHLHTDLARLGRSGGAOVSVVRSRLDPAGVATLEQIDCVRYLTHDNPGLRAAL TAADMEARAERGRIASIMGAEQGSIDNSLATLRALYALGVRYMTLTHDNNWADSA TDEPGVGLSAFGREVVRNMRGLMVLDSHVAAATMRDALDSTAPVLFSSSSRAV CDHPRNIPDDVLERLSANGAMVTFVPKFIQAAVDVTAEDDNNRAGHGFHLLDSSP EAMKHAAPFERVPRVPAVTSTVADHLDMREYVAGVDHLGIDGYDGTPTPDGLGV SGYPNLIAELLDRGWSQSLAKLTWKNVAVRVLDAEDVSRGLRAARGPSNATIEQLDG TAAAOPEG" complement(2820. .4010) /gene="SCO3058" /note="Pfam match to entry PF01244 Renal dipeptase, Renal dipeptidase, score 280.40, E-value 2.3e-80" complement(2836. .3956)							
repeat_unit									

gene /note="Degenerately, directly repeated at 12216. .13297"
complement(4025. .4567)
/gene="SCO3059"
CDS /note="synonyms: purB, SCBAC19G2.14c"
complement(4025. .4567)
/gene="SCO3059"
/EC number="4.1.1.21"
/note="SCBAC19G2.14c, phosphoribosylaminoimidazole
carboxylase catalytic subunit PurB, len: 180aa: strongly
similar to many eg. SW:Q4679 (PurB_CORAM)
phosphoribosylaminoimidazole carboxylase catalytic subunit
PurB from Corynebacterium ammoniagenes (177 aa) fasta
scores; opt: 666, Z-score: 735.8, 65.541% identity
(65.541% ungapped) in 148 aa overlap and TR:O80937
(EMBL:AC004684) putative phosphoribosylaminoimidazole
carboxylase from Arabidopsis thaliana (645 aa) fasta
scores; opt: 644, Z-score: 703.2, 60.870% identity
(60.870% ungapped) in 161 aa overlap. Contains Pfam match
to entry PF00731 AIRC, AIR carboxylase."
/codon_start=1

Query Match 46.5%; Score 734; DB 1; Length 276800;
Best Local Similarity 72.6%; Pred. No. 1.3e-128;
Matches 963; Conservative 0; Mismatches 360; Indels 3; Gaps 1;
QY 106 CGTGAATAATCTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGGAATCATG 165
DB 43666 CGTGAAGCGTCACAGTGAAGCGTCACATTTAGAAAGAGATGCTGTGCGCTCCATCGAC 43607
QY 166 CAGCTATTTCGCTCGCGAAATTCAGCTCCCGCGGTAAACCAACCGTCGAGGCGAGAGTTT 225
DB 43606 GTGCTGTAGCCGGGAAATCTTGGACTCCCGAGGCAACCCACGGTCGAGTCCGAGTGC 43547
QY 226 TTCCTGATGACGTTTCCACGGTGTGCGAGGTGTTCATCCGCGCGATCCACCGGCGTC 285
DB 43546 GGCCTCGACGACGACGAGCGGGTGTGCGCGCGCTTCCGTCCGCGCGCTCCACCGCGCC 43487
QY 286 CACGAGGCTCATGAGCTGCGTGAAGCGTGGC---GATCGCTACCTGGGCAAGGGGCTTTTG 342
DB 43486 TTCGAGGCGCATCGAGCTGCGTGAAGCGGACCGGAGCGGTACCTCGGCAAGGGGCTCGAG 43427
QY 343 AAGCAGCTTGAACCGTCAACGAAGAATCGCGACGAGCTCGCTGCGCTAGAGGCTGAC 402
DB 43426 AAGCCGCTGTGCGCGTTCATCGACAGATCGGCGGAGTGGTGGCTACGACGCCACC 43367
QY 403 GATCAGCGCTCATCGAGGAAGCAATGATCAAGCTTGATGSCACCGCGCAACAAGTCCGC 462
DB 43366 GAGCAGCGCTGATCGACCGAGCCATGTTCCACCTGGACGCCACCAACAAGGGCTCG 43307
QY 463 CTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTCGAAAGGCTGCTGATTC 522
DB 43306 CTGCGGCGCAACGCCATCTCGGGTGTCTCTCGCGTCCGCCAGCGCGCTCGAGGCC 43247
QY 523 GCAGCGCTCCCACTGTTCCGTATCATCGTGGACCAACGACACAGTCTTTCAGTTTCCA 582
DB 43246 AGCGACCTGCGCTCTTCCGTCTACCTGGCGGCGCGAAGCGCACCTGCTGCGGTGCG 43187
QY 583 ATGATGAACATCATCAAGGTTGGCGCTCACGCTGATCCCGTGTGTGAGTTTCAGGAATTC 642
DB 43186 ATGATGAACATCTTGAACCGCGGCTCGCACGCCCATCTCAACGTTGGACATCCAGGAGTTC 43127
QY 643 ATGATCGCTCAATCGGTGACAGACCTTCTCTAGGCTCTCCGCAACCGCGCGGAGGTC 702
DB 43126 ATGATCGCGCGATCGGCGGAGTCTCTTCTCCGAGGCGTGGCTGGGGGCGCGAGGTC 43067
QY 703 TACCACGACATGAAGTCCGTATCAAGAAAGAGGCTGTCTCACCGGATTTGGCGATGAG 762
DB 43066 TACCACACCTCAAGAAAGGTCCTGAAGAAACAGAGGCTTGGCACCGGCTCGGCGACGAG 43007
QY 763 GCGGCTTCGCTCTTCCGTGCGTCCACCGTGGAGGCTTTGACCTTATCTGTGAGGCA 822
DB 43006 GCGGCTTCGCGCGAACTGGGCTTCAACCGCGAGGCGCTTCGACCTCATCTCTGAGGCG 42947

QY 823 ATCAGAGAAGGCTGGCTTACCCCGAGCAAGACATCGCTTGTCTGTGAGAGCTTGTTC 882
DB 42946 ATCAAGGAGCGCGCTACACCCCGGAGCAGATCGCCCTCGGGCTCGAGCTCGCGCG 42887
QY 883 TCTGAGTTTCAAGGAGCGCACTACCACTTCAAGAGTGGCGAGCACTCCGAGCTGAG 942
DB 42886 TCCGAGTTTCAAGGAGCGCTCTACGCCCTTCAAGGAGCAAGAACCGCTCGCGCGCGAG 42827
QY 943 ATGGCAACGTTTACCGTGAAGCTGTTGAGCGGTACCAATCTCTCCATCGAGGAGCCA 1002
DB 42826 ATGACGAGTACCTACGCGGAGCTGGTTCGAGCGGTACCCGCTGGTCTCTCACTCGAGGAGCCG 42767
QY 1003 CTGCAAGAGATGACTGGGAGGCTTACACCAACTCACCCCAACCACTCGCGCGCAAGGTT 1062
DB 42766 CTGTTGAGGAGCAGCTGGGAGCGCTGGAACACCATCACCGCCAAAGCTCGGCGCAAGGTG 42707
QY 1063 CAGATCGTTGGCGAGCACTTCTTCTGTCACCAACCTGAGCGCTGAAAGGAGGCGATCGCT 1122
DB 42706 CAGCTGTGCGGAGCAGCTTCTGTCACCAACCGGAGCGCTTGGCGCGGCGATCGAG 42647
QY 1123 AAGAGGCTGCCAACTCCATCTGTTAAGGTGAACAGATCGGTACCTCCACGAGACC 1182
DB 42646 GGAAGCTCGGCCAACGCCCTGCTGCTCAAGGTCAACAGATCGGTTCGCTGACCGAGACT 42587
QY 1183 TTGAGCGCTGTGACATGGCTCACCGCGAGGCTTACACCTTCCATGATGTCCACCGTTCC 1242
DB 42586 CTGGAAGCGCTGAGCTGGCGCCAGCGCAACGGCTTCAAGTGCATGATGTCCACCGCTCC 42527
QY 1243 GGTGAGACCGAGGACCACTTGTGACCTCGAGTGTGACCTCGAGTGTGACCTAACTGTGGCGCATC 1302
DB 42526 GGGCAGACCGAGGAGCTGACATCGCGCGACCTGGCGCTGCCCACTCGCGGCGAGATC 42467
QY 1303 AAGACTGGTCTCCAGCAGCTTCCGACCGGTGCGAAAGTACAAACAGCTTCTCCGCGATC 1362
DB 42466 AAGACCGCGCGCGCGCGCGCTCCGAGCGGTGCGCAAGTACAAACAGCTGTGCTCGCGATC 42407
QY 1363 GAGCAGCTGTGCGGACGCGCGGCTTACGAGGTGCGAGCGATTTCCACGCGATTTTCAG 1422
DB 42406 GAGGAGATCTCGACGAGCGCGCGGTGTACGCGCGCGAGCGCTTCCCGCTTCAAG 42347
QY 1423 GGCTAA 1428
DB 42346 GGCTGA 42341
RESULT 15
AE016822_17/c
WPCOMMENT
Sequence split into 26 fragments LOCUS AE016822 Accession AE016822
Fragment Name Begin End
AE016822_00 1 110000
AE016822_01 100001 210000
AE016822_02 200001 310000
AE016822_03 300001 410000
AE016822_04 400001 510000
AE016822_05 500001 610000
AE016822_06 600001 710000
AE016822_07 700001 810000
AE016822_08 800001 910000
AE016822_09 900001 1010000
AE016822_10 1000001 1110000
AE016822_11 1100001 1210000
AE016822_12 1200001 1310000
AE016822_13 1300001 1410000
AE016822_14 1400001 1510000
AE016822_15 1500001 1610000
AE016822_16 1600001 1710000
AE016822_17 1700001 1810000
AE016822_18 1800001 1910000
AE016822_19 1900001 2010000
AE016822_20 200001 2110000
AE016822_21 2100001 2210000
AE016822_22 2200001 2310000

AE016822_23 2300001 2410000 AE016822_24 2400001 2510000 AE016822_25 2500001 2584158 Continuation (18 of 26) of AE016822 from base 1700001 (AE016822 Leifsonia xyli subsp. xy		Query Match 44.5%; Score 701.6; DB 1; Length 110000; Best Local Similarity 72.0%; Pred. No. 1.8e-122; Matches 930; Conservative 0; Mismatches 359; Indels 3; Gaps 1;	
Qy	140	AGGAGGCGACAGTGGCTGAATATCATGACGATTCGCTCGCGAATTTCTCGACTCCCGCG	199
Db	83109	AGGAGACTGTGTGGCTGCAATCGAAGCTGTAGGAGCTCGGAGATCTCGACTCGCGCG	83050
Qy	200	GTAACCCCAACCGTCGAGGACAGAGTTCCTCGGATGACGGTTCCCAACGGTGTGCGAGGTG	259
Db	83049	GAACCCCGACCGTCGAGGTGCGAGTGTCTCTGAGATGGACGGTCAAGCGCGCGCGCG	82990
Qy	260	TTCCATCCGGCGCATCCACCGCGTCGACGAGGCTCATGAGTGGTGTGACGGTGGC---G	316
Db	82989	TCCGTCGCGGCGGTCCACCGCGCGTTCGAGGCTTACGAGTCCGCGACGCGCGACAAGG	82930
Qy	317	ATCGCTACTCGGCAAGGGCGTTTGAAGGCGAGTTGAAACGTCAACGAAGAAATCGGCG	376
Db	82929	GCCGCTACTGCGCAAGGGGTTCGAGAAGGCGGTGGATGCCGTGCTCGACGAGATCGGCC	82870
Qy	377	ACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAAGAAATGATCAAGC	436
Db	82869	CGGCCATCGAGGCTTCGAGGCCAGCGACAGCGCATCGTCGATGAGCGGATGATCGAAC	82810
Qy	437	TGATGGACCGCCCAAGTCCCGCTCGGTCGAAACGCAATCTTGGTGTTCATGG	496
Db	82809	TCGACGGCACCGCAACAAAGAAAGCGCTGGGCGCTTAAGCCATGCTCGGTGTGAGCCTCG	82750
Qy	497	CTGTGTGAAAGCTGCTGCTGATTCGCGAGCGCTCCCACTGTCGCTACATCGGTGAC	556
Db	82749	CAGTCGCAAGGCGCCCGGCTCCGCGACCTGCGCGCTGTTCGCTTACCTGGGCGGCC	82690
Qy	557	CAAAACGCAACAGTTCTTCAGTTCCAAATGATGAACATCAATCAAGGTCGCGCTCACCGTG	616
Db	82689	CGAACGGCGATGTCTCTGCGCGCTCGGATGATGAACATCAACGGCGCGCGCACCGCG	82630
Qy	617	ACTCGGTGTGAGCTTCAGGAAATTCATGATGCTGCTCAATCGGTGCGAGACCTTCTCTG	676
Db	82629	ACACCGGCTCGACATCCAGGAGTTCAATTCCTCCGATCGGCGCGGAGACCTTCTCCG	82570
Qy	677	AGGCTCTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTCCGTATCAAGGAAAGG	736
Db	82569	AGGGCTCGCTGGGGCGGTGAGAGCTTACCACTCGCTCAAGGCGCTGCTGAAGTCGAAGG	82510
Qy	737	GCCTGTCCACCGGACTTGGCGATGAGGCGCTTCCGCTCCTTCCGTCGGCTCCACCGGTG	796
Db	82509	GCCTGAACACCGGCTGGCGACGAGGCGGCTTCGCCCGCGAGCTGAGGACACACCGCG	82450
Qy	797	AGGCTCTTGAACCTTATCGTTGAGGCAATCGAGAAGGCTGGCTTACCCCGAGGCAAGGACA	856
Db	82449	CTGGCTCGACCTGATCGCCGAGGCCATCGAGAAGGCGGCTTCAACGTGCGCTCGCAGA	82390
Qy	857	TCGCTCTGTGCTGTGACGCTTCTCTGAGTTCCTCAAGGACGGCACCTTACCACTTCG	916
Db	82389	TCGGCTCGGCTCGAOCGTGCGCTCCACCGAGTTCTTCGAGAACGGCGCTTACCGCTTCG	82330
Qy	917	AAGGTGGCGACGACTCCGCGATGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGT	976
Db	82329	AGGGCGAGGACCGCACCGCGCGCGGATGAGGCGCTACTACCGCATCTCGGAACAACT	82270
Qy	977	ACCCAATCGTCTCCATCGAGGACCCACTGACGAGGAGATGACTGGAGGGTTTACACCAACC	1036
Db	82269	ACCGCTGTCTCCATCGAGGACCGGCTGGCGGAGGACGACTGGGAGGGCTGGGCGCACCC	82210
Qy	1037	TCACGGCAACCATCGGCGAAGAGTTGATCGTTGGGAGGACTTCTTGTGTCACCAACC	1096
Db	82209	TGAACGCGGAGATCGGCTCGACGCTGCAACTCGTCGGCGATGACCTGTTCTGTCACCAACC	82150

Search completed: September 28, 2005, 20:45:20
 Job time : 6718 secs

Qy	1097	CTGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCCAATCCATCCTGTTAAGGTGA	1156
Db	82149	CGAAGCGCTTGGCGCAGGGCATCCAGCAGAAAGGCCGGAACAGCATCTGCTCAAGGTGA	82090
Qy	1157	ACGAGATCGGTACCTTACCGAGACCTTTCGACGCTGTGACATGGCTACCGCGAGGCT	1216
Db	82089	ACCAGATCGGCACCTTTCGACGAGACGCTGGACGCGGTCTCGCTCGCCCGAGCGCGTA	82030
Qy	1217	ACACTCCATGATGTCCACCGTTCCGTTGAGACCGAGGACCAACCATTCCTGACCTCG	1276
Db	82029	TGACCGCGCTCTCTCGACCGCTCCGCGGAGACCGAGGACCAACCATTCGCGACCTCG	81970
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTGCTCCAGCACGTTCCGACCGTGTG	1336
Db	81969	CCGTGCGCACCGATTTCGCGCCAGATCAAGACCGCGCGCCCGGCGCTCCGAGCGCGTCG	81910
Qy	1337	CAAAAGTACAACCAAGCTTCTCGGATCGAGCAGCTGCTTGGGAGCGCGGCGTCTTACGAG	1396
Db	81909	CGAAGTACAACCAAGCTTCTCGGATCGAGGAGGAGCTGGGCGAGGCGGCTGTCTACGCG	81850
Qy	1397	GTGCGACGCAATTCACCGCTTTCAGGGCTAA	1428
Db	81849	CCCCTTCGGCGTTCCCCCGTTTACCAGGGCTGA	81818

THIS PAGE BLANK (user)

THIS PAGE BLANK (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:00:03 ; Search time 793.5 Seconds
(without alignments)
11772.340 Million cell updates/sec

Title: US-10-728-947-1

Perfect score: 1578

Sequence: 1 ggctgggatatgggtagt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	4	Aaf61696 C. glutam
2	1578	100.0	349980	5	Aah68527 C. glutami
3	1403.4	88.9	1405	10	Add13322 C. glutami
4	1396.4	88.5	1398	4	Aaf71395 Coryneb
5	1278	81.0	1278	8	Acc45311 Coryneb
6	1275	80.8	1275	5	Aah6050 C. glutami
7	1275	80.8	1299	8	Acc45312 Coryneb
8	1273.2	80.7	1278	8	Acc45314 Modified
9	1270.2	80.5	1299	8	Acc45313 Modified
10	908.6	57.6	1275	8	Acc29854 Prokaryot
11	679.8	43.1	1269	13	Ads56353 Bacterial
12	647	41.0	1287	8	Acc38079 Prokaryot
13	642.4	40.7	30159	4	Aas59564 Propionib
14	642.4	40.7	30159	4	Acc64493 Propionib
15	632.6	40.1	45190	6	Abx09144 Mycobacte
16	632.6	40.1	110000	4	Continuation (12 o
17	632.6	40.1	110000	4	Continuation (12 o
18	632.4	40.1	1317	8	Acc38642 Prokaryot
19	626.2	39.7	1290	8	Acc40436 Prokaryot
20	584.6	37.0	349980	6	Abq81846 Bifidobac

c	21	584.6	37.0	349980	6	Abq81847
	22	583.4	37.0	1344	8	ACA39628
	23	582	36.9	1269	13	ADS58570
	24	575.4	36.5	1269	13	ADS59709
	25	574	36.4	1275	13	ADS59832
	26	574	36.4	1281	13	ADT42160
	27	569.8	36.1	1278	13	ADT43135
	28	569	36.1	1272	13	ADT45911
	29	567.4	36.0	1260	13	ADS62752
	30	567.4	36.0	1269	13	ADS63235
	31	561.8	35.6	1272	13	ADS55775
	32	557	35.3	1278	13	ADT46979
	33	552	35.0	1290	13	ADT44740
	34	542	34.3	1281	8	ACA23750
	35	533.6	33.8	1335	11	ABD05845
	36	533	33.8	1347	11	ABD05554
	37	531.8	33.7	1290	4	AAS54197
	38	531.8	33.7	1290	8	ACA42456
	39	530.8	33.6	1281	13	ADT46255
	40	528	33.5	1287	13	ADS56507
	41	522.8	33.1	1278	13	ADS49945
	42	522.8	33.1	1281	8	ACA25228
	43	522.8	33.1	1281	13	ADS55460
	44	517.4	32.8	1690	12	ADO59752
	45	515.2	32.6	1404	8	ACA45728

ALIGNMENTS

RESULT 1

AAF61696
ID AAF61696 standard; DNA; 1578 BP.

XX AAF61696;

DT 12-JUL-2001 (first entry)

DE C. glutamicum enolase encoding DNA.

XX Enolase; fermentation; L-amino acid; L-lysine; coryneform; eno gene;
KW Enolase; animal feed supplement; ds.

OS Corynebacterium glutamicum.

PH Key Location/Qualifiers

FT CDS 151..1428

FT /*tag= a

FT /product= "Enolase"

XX EPI090998-A1.

XX PD 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121158.

XX 05-OCT-1999; 99DE-01047791.

XX (DEGS) DEGUSSA-HUELS AG.

XX Moekkel B, Pfeifferle W, Hermann T, Puehler A, Kalinowski J;

XX Bathe B;

XX WPI; 2001-292928/31.

XX P-PSDB; AAB70882.

XX New enolase gene from coryneform bacteria, used to prepare transformants

XX with increased synthesis of amino acids, particularly lysine.

XX Claim 4; Page 12-14; 25pp; German.

XX This invention describes a novel isolated nucleic acid (I) from

XX coryneform bacteria which is used in a method for fermentative production

PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX Disclosure; SEQ ID NO 7062; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of coryneform bacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Coryneform bacterium, and identifying a homologue of a gene derived from
XX coryneform bacterium. Coryneform bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the European Patent Office
XX
SQ Sequence 349980 BP; 79725 A; 90426 C; 98918 G; 80911 T; 0 U; 0 Other;
Query Match 100.0%; Score 1578; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCTGGGGATATGGGTAGTGTTCGCCACTAATTTCACTGATTCGCTCATCGAAACAAGA 60
Db 134799 GGCTGGGGATATGGGTAGTGTTCGCCACTAATTTCACTGATTCGCTCATCGAAACAAGA 134858
Qy 61 TTCGTGCACAAATTTGGGTGTAGAGTGATTTGAAGACATTTGATCAGGTGAATAATTTCTAG 120
Db 134859 TTCGTGCACAAATTTGGGTGTAGAGTGATTTGAAGACATTTGATCAGGTGAATAATTTCTAG 134918
Qy 121 TTAGTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCAGTATTCGCTCGC 180
Db 134919 TTAGTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCAGTATTCGCTCGC 134978
Qy 181 GAAATTTCTCGACTCCCGCGTAAACCCACCGTCCGAGGACAGAGTTTTCCTGGATGACGGT 240
Db 134979 GAAATTTCTCGACTCCCGCGTAAACCCACCGTCCGAGGACAGAGTTTTCCTGGATGACGGT 135038
Qy 241 TCCACCGGTGTCCAGAGTGTTCATTCGCGCGCATCCACCGCGTCCAGAGGCTCATGAG 300
Db 135039 TCCACCGGTGTCCAGAGTGTTCATTCGCGCGCATCCACCGCGTCCAGAGGCTCATGAG 135098
Qy 301 CTGCGTGCAGTGGCGATCGCTACTCTGGCAAGGCGGTTTTGAAGGCGATTTGAAGACGTC 360
Db 135099 CTGCGTGCAGTGGCGATCGCTACTCTGGCAAGGCGGTTTTGAAGGCGATTTGAAGACGTC 135158
Qy 361 AACGAAGAAATCGGCGACGAGTGGCTGGCCCTAGAGGCTGACGATCAGCGGCTCATCGAC 420
Db 135159 AACGAAGAAATCGGCGACGAGTGGCTGGCCCTAGAGGCTGACGATCAGCGGCTCATCGAC 135218
Qy 421 GAAGCAATGATCAAGCTTGTATGGCAACCGCAACAGTCCCGCTGGGTGCAAAACGCAATC 480
Db 135219 GAAGCAATGATCAAGCTTGTATGGCAACCGCAACAGTCCCGCTGGGTGCAAAACGCAATC 135278
Qy 481 CTTGGTGTTCATGGCTGTTGCAAGAGCTGCTGATTCGGCAGGCTCCCACTGTTTC 540
Db 135279 CTTGGTGTTCATGGCTGTTGCAAGAGCTGCTGATTCGGCAGGCTCCCACTGTTTC 135338
Qy 541 CGCTACATCGGTGGACCAACCGCACAGTCTTCTCCAGTTTCCAAATGATGAACATCATCAAC 600

Db 135339 CGCTACATCGGTGGACCAACCGCACAGTCTTCTCCAGTTCCAATGATGAACATCATCAAC 135398
Qy 601 GGTGGCGCTCAGCTGACTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCAAATCGGT 660
Db 135399 GGTGGCGCTCAGCTGACTCGGTGTTGAGCTTCAGGAATTCATGATCGCTCCAAATCGGT 135458
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACACGCACTGAAGTCC 720
Db 135459 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACACGCACTGAAGTCC 135518
Qy 721 GTCATCAAGAAAAAGGCCCTGTCCACCGCATTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
Db 135519 GTCATCAAGAAAAAGGCCCTGTCCACCGCATTTGGCGATGAGGCGGCTTCGCTCCTTCC 135578
Qy 781 GTGGGCTCCACCCGTGAGGCTCTTGACCTTATCGTTGAGCAATCGAGAGGCTGGCTTC 840
Db 135579 GTGGGCTCCACCCGTGAGGCTCTTGACCTTATCGTTGAGCAATCGAGAGGCTGGCTTC 135638
Qy 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTTCTCTCTGATTTCTTCAAGGAC 900
Db 135639 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTTCTCTCTGATTTCTTCAAGGAC 135698
Qy 901 GGCACTTACCACTTCGAAGGTGGCCAGACTCCGCACTGAGATGGCAAAAGCTTTACGCT 960
Db 135699 GGCACTTACCACTTCGAAGGTGGCCAGACTCCGCACTGAGATGGCAAAAGCTTTACGCT 135758
Qy 961 GAGTCGTTGACGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAGATGACTGG 1020
Db 135759 GAGTCGTTGACGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAGATGACTGG 135818
Qy 1021 GAGGTTTACACCACTTCACCGCAACCATCGGCAAGGTTTCAGATCGTTGGCGACGAC 1080
Db 135819 GAGGTTTACACCACTTCACCGCAACCATCGGCAAGGTTTCAGATCGTTGGCGACGAC 135878
Qy 1081 TTCTTCGTCAACCACTGAGCGCTGAAGAGGGCATTCGTAAAGAGGCTTGCCAACTCC 1140
Db 135879 TTCTTCGTCAACCACTGAGCGCTGAAGAGGGCATTCGTAAAGAGGCTTGCCAACTCC 135938
Qy 1141 ATCTCGGTTAAGGTGAACCACTGATACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db 135939 ATCTCGGTTAAGGTGAACCACTGATACCTCACCGAGACCTTCGACGCTGTCGACATG 135998
Qy 1201 GCTCACCGCGCAGGCTACACCTCATGATGTCCCAACCGTTCCGGTGAGACCGAGGACAC 1260
Db 135999 GCTCACCGCGCAGGCTACACCTCATGATGTCCCAACCGTTCCGGTGAGACCGAGGACAC 136058
Qy 1261 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA 1320
Db 136059 ACCATTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA 136118
Qy 1321 GGTTCGACCGTGTGCGCAAAAGTACAAACAGGTTCTCCGCATCGAGGAGCTGCTTGGCGAC 1380
Db 136119 GGTTCGACCGTGTGCGCAAAAGTACAAACAGGTTCTCCGCATCGAGGAGCTGCTTGGCGAC 136178
Qy 1381 GCGCGGCTGTACGCGAGGTGCGCAGGCAATCCACGCTTTCAGGGCTAAATAAAGCGCTT 1440
Db 136179 GCGCGGCTGTACGCGAGGTGCGCAGGCAATCCACGCTTTCAGGGCTAAATAAAGCGCTT 136238
Qy 1441 TTCGACCGCCCGGTAACTCAAGGTTGCGGCGCTGTTGCTTACTACTGTACTGTGTGT 1500
Db 136239 TTCGACCGCCCGGTAACTCAAGGTTGCGGCGCTGTTGCTTACTACTGTACTGTGTGT 136298
Qy 1501 GACTATGATCGAGGATTTATGGCAAGGAGGAGAACTCATAAAGGCTTGTCTCTGTCT 1560
Db 136299 GACTATGATCGAGGATTTATGGCAAGGAGGAGAACTCATAAAGGCTTGTCTCTGTCT 136358
Qy 1561 CAACGAGGAACTGCTT 1578
Db 136359 CAACGAGGAACTGCTT 136376
RESULT 3

Qy 1191 TGTGACATGCTCACCGCGAGGCTACACCTTCATGATGTCCACCGTTCGGTGAGAC 1250
 Db 1141 TGTGACATGCTCACCGCGAGGCTACACCTTCATGATGTCCACCGTTCGGTGAGAC 1200
 Qy 1251 CGAGGACACCACTTGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCACTTGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
 Qy 1311 TGCTCCAGCAGTTCCGACCGTGTGCGAAAGTAAACACAGCTTCTCCGATCGAGCAGCT 1370
 Db 1261 TGCTCCAGCAGTTCCGACCGTGTGCGAAAGTAAACACAGCTTCTCCGATCGAGCAGCT 1320
 Qy 1371 GCTTGGGACCGCGCTTACGAGGTGCGAGGTGCGAGCAGTTCACAGCTTTCAGGCTAAAT 1430
 Db 1321 GCTTGGGACCGCGCTTACGAGGTGCGAGGTGCGAGCAGTTCACAGCTTTCAGGCTAAAT 1380
 Qy 1431 AAAAGCGCTTTTCGACGCCCGGTAA 1455
 Db 1381 AAAAGCGCTTTTCGACGCCCGGTAA 1405

RESULT 4

ID AAF71395 standard; DNA; 1398 BP.

XX AAF71395;

XX 30-APR-2001 (first entry)

DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:71.

XX Corynebacterium glutamicum; carbon metabolism and energy production;
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.

XX Corynebacterium glutamicum.

XX WO200100844-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB000943.

XX 25-JUN-1999; 99US-0141031P.

PR 08-JUL-1999; 99DE-01031412.

PR 08-JUL-1999; 99DE-01031413.

PR 08-JUL-1999; 99DE-01031419.

PR 08-JUL-1999; 99DE-01031420.

PR 08-JUL-1999; 99DE-01031424.

PR 08-JUL-1999; 99DE-01031428.

PR 08-JUL-1999; 99DE-01031431.

PR 08-JUL-1999; 99DE-01031433.

PR 08-JUL-1999; 99DE-01031434.

PR 08-JUL-1999; 99DE-01031510.

PR 08-JUL-1999; 99DE-01031562.

PR 08-JUL-1999; 99DE-01031634.

PR 09-JUL-1999; 99DE-01032180.

PR 09-JUL-1999; 99DE-01032227.

PR 09-JUL-1999; 99DE-01032230.

PR 03-SEP-1999; 99DE-01042088.

PR 03-SEP-1999; 99DE-01042095.

PR 03-SEP-1999; 99DE-01042123.

PR 03-SEP-1999; 99DE-01042125.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habershauer G;

PI WPI; 2001-061975/07.

DR P-PSDB; AAB79278.

XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar

PT metabolism and oxidative phosphorylation protein for production or

PT modulation of production of fine chemicals e.g. amino acids,

PT carbohydrates or enzymes.

XX Claim 3; Page 242-244; 1246pp; English.

XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localise C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)

XX SQ Sequence 1398 BP; 308 A; 423 C; 376 G; 291 T; 0 U; 0 Other;

XX Query Match 88.5%; Score 1396.4; DB 4; Length 1398;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 1397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 51 CGAAACAAGATTTCGTGCAACAATTGGGTGTAGAGCTGATTGAAGACATTTTCATCAGTGA 110

Db 1 CGAAACAAGATTTCGTGCAACAATTGGGTGTAGAGCTGATTGAAGACATTTTCATCAGTGA 60

Qy 111 ATAACTCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGT 170

Db 61 ATAACTCTAGTTAGTCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGT 120

Qy 171 ATTGCTCGGGAATTTCTCGACTCCCGGGTAAACCAACCGTTCGAGGACAGAGTTTTCCT 230

Db 121 ATTGCTCGGGAATTTCTCGACTCCCGGGTAAACCAACCGTTCGAGGACAGAGTTTTCCT 180

Qy 231 GGATGACGGTTCCACCGGTGTCGAGGTGTTCCATCGGGCGCATCCACCGCGTCCACGA 290

Db 181 GGATGACGGTTCCACCGGTGTCGAGGTGTTCCATCGGGCGCATCCACCGCGTCCACGA 240

Qy 291 GGCTCATGAGCTCGTGACGGTGGCGATCGCTACTCGGGCAAGGGCGTTTTGAAGGCAGT 350

Db 241 GGCTCATGAGCTCGTGACGGTGGCGATCGCTACTCGGGCAAGGGCGTTTTGAAGGCAGT 300

Qy 351 TGAACAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTAGAGGCTAGCATCAGCG 410

Db 301 TGAACAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTAGAGGCTAGCATCAGCG 360

Qy 411 CCTCATCGAGCAATGATCATCAAGCTTGATGGCACCGCCAAAGTCCCGCTGGGTGC 470

Db 361 CCTCATCGAGCAATGATCATCAAGCTTGATGGCACCGCCAAAGTCCCGCTGGGTGC 420

Qy 471 AACGCAATCCTTGGTGTTCCTCATGGGTGTTTGCAGAGGCTGCTGCTGATTCGCGAGGCCT 530

Qy 331 AAGGCGGTTTGAAGCAGTTGAAAACTCAACGAAGAAATCGGCGACGAGCTCGCTGGC 390
 Db |||||
 Qy 181 AAGGCGGTTTGAAGCAGTTGAAAACTCAACGAAGAAATCGGCGACGAGCTCGCTGGC 240
 Db |||||
 Qy 391 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGACCGCC 450
 Db |||||
 Qy 241 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGACCGCC 300
 Db |||||
 Qy 451 AACAACTCCGCTGGGTGCAACGCAATCTCTGTTTCCATGGCTGTTGCAAGGCT 510
 Db |||||
 Qy 301 AACAACTCCGCTGGGTGCAACGCAATCTCTGTTTCCATGGCTGTTGCAAGGCT 360
 Db |||||
 Qy 511 GCTGCTGATTCCGACGAGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 570
 Db |||||
 Qy 361 GCTGCTGATTCCGACGAGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 420
 Db |||||
 Qy 571 CTTCCAGTTCGAATGATGAACATCATCAACGCTGGCTCAGCTGACTCCGGTGTGAC 630
 Db |||||
 Qy 421 CTTCCAGTTCGAATGATGAACATCATCAACGCTGGCTCAGCTGACTCCGGTGTGAC 480
 Db |||||
 Qy 631 GTTCAGGAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
 Db |||||
 Qy 481 GTTCAGGAATTCATGATCGCTCCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 540
 Db |||||
 Qy 691 GGCACGAGGCTTACCAACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGA 750
 Db |||||
 Qy 541 GGCACGAGGCTTACCAACGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGA 600
 Db |||||
 Qy 751 CTTGGCGATGAGGCGGCTTCCGCTCCTTCCGCTGGCTCCACCCGTGAGGCTCTTGACCTT 810
 Db |||||
 Qy 601 CTTGGCGATGAGGCGGCTTCCGCTCCTTCCGCTGGCTCCACCCGTGAGGCTCTTGACCTT 660
 Db |||||
 Qy 811 ATCGTTGAGGCAATCGAGAGGCTGCTTCAACCCAGGAGGATCGCTCTGCTCTG 870
 Db |||||
 Qy 661 ATCGTTGAGGCAATCGAGAGGCTGCTTCAACCCAGGAGGATCGCTCTGCTCTG 720
 Db |||||
 Qy 871 GAGCTTCTCTCTGAGTCTTCAAGGAACGCACTTACCACTTCGAAGGTGGCCAGCAC 930
 Db |||||
 Qy 721 GAGCTTCTCTCTGAGTCTTCAAGGAACGCACTTACCACTTCGAAGGTGGCCAGCAC 780
 Db |||||
 Qy 931 TCCGACGCTGAGTGGCAAGCTTTAGCTGAGTCTGTTGACGCTACCAATCGTCTCC 990
 Db |||||
 Qy 781 TCCGACGCTGAGTGGCAAGCTTTAGCTGAGTCTGTTGACGCTACCAATCGTCTCC 840
 Db |||||
 Qy 991 ATCGAGGACCACTGACGAGGATGATCGGAGGTTTACACCACTTCAACCGCAACATC 1050
 Db |||||
 Qy 841 ATCGAGGACCACTGACGAGGATGATCGGAGGTTTACACCACTTCAACCGCAACATC 900
 Db |||||
 Qy 1051 GGCACAAAGGTTCAAGATCGTTGGCGAGGACTTCTTCTGTCACCAACCTGAGCGCTGAAG 1110
 Db |||||
 Qy 901 GGCACAAAGTTCAGATCGTTGGCGAGGACTTCTTCTGTCACCAACCTGAGCGCTGAAG 960
 Db |||||
 Qy 1111 GAGGCGATCGTAAGAGGCTGCCAATCCATCTGTTTAAAGTGAACCAATCGGTACC 1170
 Db |||||
 Qy 961 GAGGCGATCGTAAGAGGCTGCCAATCCATCTGTTTAAAGTGAACCAATCGGTACC 1020
 Db |||||
 Qy 1171 CTCACCGAGACCTTCGAGCTGTGACATCGCTCACCAGGAGGCTACCTCCATCATG 1230
 Db |||||
 Qy 1021 CTCACCGAGACCTTCGAGCTGTGACATCGCTCACCAGGAGGCTACCTCCATCATG 1080
 Db |||||
 Qy 1231 TCCACCGCTTCGCTGAGACCGAGGACCAACCAATTCGTACCTTCGAGTGTGACTCAAC 1290
 Db |||||
 Qy 1081 TCCACCGCTTCGCTGAGACCGAGGACCAACCAATTCGTACCTTCGAGTGTGACTCAAC 1140
 Db |||||
 Qy 1291 TGTGGCCAGATCAAGACTGTGTCCAGCAGGTTCCGACCGTGTCCGAAAGTACCAACAG 1350
 Db |||||
 Qy 1141 TGTGGCCAGATCAAGACTGTGTCCAGCAGGTTCCGACCGTGTCCGAAAGTACCAACAG 1200
 Db |||||
 Qy 1351 CTTTCTCCGATCGAGCAGCTGCTTGGCGACCGCGGCTCTACGAGGCTCGGAGGCGATT 1410
 Db |||||
 Qy 1201 CTTTCTCCGATCGAGCAGCTGCTTGGCGACCGCGGCTCTACGAGGCTCGGAGGCGATT 1260
 Db |||||

Qy 1411 CCACGCTTTCAGGGCTAA 1428
 Db |||||
 Qy 1261 CCACGCTTTCAGGGCTAA 1278
 Db |||||

RESULT 6

AAH66050
 ID AAH66050 standard; DNA; 1275 BP.
 XX AC AAH66050;
 XX AC
 DT 26-SEP-2001 (first entry)
 XX XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 1085.
 XX XX
 KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX XX
 OS Corynebacterium glutamicum.
 XX XX
 PN EPI108790-A2.
 XX XX
 PD 20-JUN-2001.
 XX XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayaashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX XX
 DR WPI; 2001-376931/40.
 DR P-PSDB; AAG90831.

Novel polynucleotides derived from Corynebacterium bacteria, for identifying mutation point of a gene, measuring expression of a gene, analyzing expression profile or pattern of a gene and identifying homologous gene.
 Claim 8; SEQ ID NO 1085; 246pp + Sequence Listing; English.
 The present invention provides a number of nucleotide and protein sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These are useful for identifying the mutation point of a gene derived from a mutant of corynebacterium bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from Corynebacterium bacterium, and identifying a homologue of a gene derived from Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

Seq Sequence 1275 BP; 269 A; 399 C; 347 G; 260 T; 0 U; 0 Other;

Query Match 80.8%; Score 1275; DB 5; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 3.6e-311;
 Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 GTGGCTGAATCATGACGATTCGCTCGGAAATTCGCACTCCCGGGTAACCCAAAC 210
 Db 1 GTGGCTGAATCATGACGATTCGCTCGGAAATTCGCACTCCCGGGTAACCCAAAC 60
 Qy 211 GTGAGGACAGAGGTTTCTTGGATGACGCTTCCACCGTGTCCGAGGTGTTCCATCCGGC 270
 Db 61 GTGAGGACAGAGGTTTCTTGGATGACGCTTCCACCGTGTCCGAGGTGTTCCATCCGGC 120
 Qy 271 GCATCCACCGGCTCCACGAGGCTCATGAGCTCGGTGACGCTGGCGATCGCTACCTCGGC 330
 Db |||||

Db 121 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCTGAGCGTGGCGATCGCTACCTGGGC 180
 Qy 331 AAGGCGCTTTTGAAGGCGAGTTGAAACGCTCAAAGAAATCGCGAGAGAGCTCGCTGGC 330
 Db 181 AAGGCGCTTTTGAAGGCGAGTTGAAACGCTCAAAGAAATCGCGAGAGAGCTCGCTGGC 240
 Qy 391 CTAGAGGCTCAGCATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGTATGCGACCGCC 450
 Db 241 CTAGAGGCTCAGCATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGTATGCGACCGCC 300
 Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGCAAGGCT 510
 Db 301 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGCAAGGCT 360
 Qy 511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGTACATCGTGGACCAAAAGCAAGCTT 570
 Db 361 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGTACATCGTGGACCAAAAGCAAGCTT 420
 Qy 571 CTTCCAGTTCGAATGATGAACATCATCAACGGTGGCGCTCAGCGTGAATCCGGTGTGAC 630
 Db 421 CTTCCAGTTCGAATGATGAACATCATCAACGGTGGCGCTCAGCGTGAATCCGGTGTGAC 480
 Qy 631 GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
 Db 481 GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 540
 Qy 691 GGCGCGAGGTCTACCAAGCACTGAAGTCCGTGATCAAGGAAAGGGCGCTGTCCACCGGA 750
 Db 541 GGCGCGAGGTCTACCAAGCACTGAAGTCCGTGATCAAGGAAAGGGCGCTGTCCACCGGA 600
 Qy 751 CTTGCGCATGAGGCGGCTTCCGCTTCCGTGGCTCCACCGTGAAGGCTCTTGACCTT 810
 Db 601 CTTGCGCATGAGGCGGCTTCCGCTTCCGTGGCTCCACCGTGAAGGCTCTTGACCTT 660
 Qy 811 ATCGTTGAGGCAATCGAAGGCTGGCTTACCCAGGCAAGGACATCGCTTCTGCTCTG 870
 Db 661 ATCGTTGAGGCAATCGAAGGCTGGCTTACCCAGGCAAGGACATCGCTTCTGCTCTG 720
 Qy 871 GACGTTGCTTCTGAGTCTTCAAGGACGGACCTTACCACTTCAAGGTTGGCGACGAC 930
 Db 721 GACGTTGCTTCTGAGTCTTCAAGGACGGACCTTACCACTTCAAGGTTGGCGACGAC 780
 Qy 931 TCCGACGTGAGTGGCAAGCTTTAGCTGAGCTGTTGACGGGTACCCCAATCGTCTCC 990
 Db 781 TCCGACGTGAGTGGCAAGCTTTAGCTGAGCTGTTGACGGGTACCCCAATCGTCTCC 840
 Qy 991 ATCGAGGACCCACTGCAGGAAGATGACTGGAGGGTTACCAACCTCACCGCAACCATC 1050
 Db 841 ATCGAGGACCCACTGCAGGAAGATGACTGGAGGGTTACCAACCTCACCGCAACCATC 900
 Qy 1051 GCGCAAGGTTGATGCTGGGCGACGACTTCTGTCACCAACCTGAGCGCTGAAG 1110
 Db 901 GCGCAAGGTTGATGCTGGGCGACGACTTCTGTCACCAACCTGAGCGCTGAAG 960
 Qy 1111 GAGGCACTGCTAAGAGGCTGCCAATCTCCTGTTAAGTGAACAGATCGGTACC 1170
 Db 961 GAGGCACTGCTAAGAGGCTGCCAATCTCCTGTTAAGTGAACAGATCGGTACC 1020
 Qy 1171 CTCACCGAGACTTCGACGCTGTGACATGCTCACCGCGAGGCTACACCTCATGATG 1230
 Db 1021 CTCACCGAGACTTCGACGCTGTGACATGCTCACCGCGAGGCTACACCTCATGATG 1080
 Qy 1231 TCCACCGTTCGGTGAGACCGAGACACACCATTCGTGACCTCGCATGTCACCTCAAC 1290
 Db 1081 TCCACCGTTCGGTGAGACCGAGACACACCATTCGTGACCTCGCATGTCACCTCAAC 1140
 Qy 1291 TGTGCGCAGATCAAGACTGGTGTCCAGACGTTCCGACCGTGTCCGAAAGTACAACAG 1350
 Db 1141 TGTGCGCAGATCAAGACTGGTGTCCAGACGTTCCGACCGTGTCCGAAAGTACAACAG 1200
 Qy 1351 CTTTCCGACGAGCAGCTGTTGGCGACCGCGCGTCTACGAGGTCGAGCGCATTC 1410
 Db 1201 CTTTCCGACGAGCAGCTGTTGGCGACCGCGCGTCTACGAGGTCGAGCGCATTC 1260

Qy 1411 CCACGCTTTTCAGGC 1425
 Db 1261 CCACGCTTTTCAGGC 1275

RESULT 7

ACC45312
 ID ACC45312 standard; DNA; 1299 BP.

XX ACC45312;

DT 17-JUN-2003 (first entry)

DE Corynebacterium glutamicum eno DNA fragment SEQ ID NO:3.

XX Fine chemical; Corynebacterium; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno; gene; ds.

XX Corynebacterium glutamicum.

OS WO2003023016-A2.

PN 20-MAR-2003.

XX 11-SEP-2002; 2002WO-EP010174.

XX 13-SEP-2001; 2001DE-01045043.

XX (DEGS) DEGUSSA AG.

XX Farwick M, Hermann T;

XX WPI; 2003-354534/33.

Microorganism useful for producing e.g. fine chemicals, has permanently altered phosphorylatability protein, such that biosynthesis of fine chemical synthesized by microorganism is increased compared to wild-type.

Example 3; Page 81-82; 120pp; English.

The present invention describes a microorganism (I), in which the phosphorylatability of at least one protein has been permanently altered such that the biosynthesis of at least one fine chemical synthesised by the microorganism is increased compared to the wild type. Also described: (1) use of a DNA (II) sequence coding for a protein which contains a phosphorylation site, where the sequence contains such a mutation that the protein is changed in its phosphorylatability for the production of (1), or for the production of fine chemicals; and (2) a method for producing fine chemicals or metabolites comprising using (1). (I) is useful for producing fine chemicals or metabolites, such as amino acids, vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids and vitamins produced using (I) can be used in human medicine, in the pharmaceutical industry, food industry and in animal feeding. (I) produces larger amount of desired fine chemical or a metabolite than the wild type. The present sequence represents a wild type eno fragment DNA sequence from Corynebacterium glutamicum, which is used in an example from the present invention

Sequence 1299 BP; 274 A; 409 C; 353 G; 263 T; 0 U; 0 Other;

Query Match 80.8%; Score 1275; DB 8; Length 1299;
 Best Local Similarity 100.0%; Pred. No. 3.6e-311;
 Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 154 GCTGAATCATGACGATTCGCTCGGAAATTCGACTCCCGCGTAAACCCACCGTC 213
 Db 10 GCTGAATCATGACGATTCGCTCGGAAATTCGACTCCCGCGTAAACCCACCGTC 69
 Qy 214 GAGCAGAGGTTTCTCTGGATGACGGTCCACGGTTCGAGGTTCATCCATCGCGGCA 273

Db 70 GAGGAGAGGTTTCTCGATGACGGTTCCACGGGTGTCGAGGTGTTCATCCGGGCA 129
Qy 274 TCCACCGGCTCCACAGGCTCATGAGCTGGTGAACGGTGGCGATCGTACTCGGCAAG 333
Db 130 TCCACCGGCTCCACAGGCTCATGAGCTGGTGAACGGTGGCGATCGTACTCGGCAAG 189
Qy 334 GCGCTTTTGAAGGAGTTGAAAAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTA 393
Db 190 GCGCTTTTGAAGGAGTTGAAAAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGGCTA 249
Qy 394 GAGCTGACGATCAGCGCTCATCGAGCAAGCAATGATCAAGCTTGATGGCACCGCAAC 453
Db 250 GAGCTGACGATCAGCGCTCATCGAGCAAGCAATGATCAAGCTTGATGGCACCGCAAC 309
Qy 454 AAGTCCCGCTGGGTGCAAAACGCAATCCTTGGTGTTCATGGCTGTTCGCAAGGCTGCT 513
Db 310 AAGTCCCGCTGGGTGCAAAACGCAATCCTTGGTGTTCATGGCTGTTCGCAAGGCTGCT 369
Qy 514 GCTGATTCGGAGGCTCCCACTGTTTCGGCTACATCGGTGGAACCAAGCAACAGTCTTT 573
Db 370 GCTGATTCGGAGGCTCCCACTGTTTCGGCTACATCGGTGGAACCAAGCAACAGTCTTT 429
Qy 574 CCAGTCCCAATGATGAACATCATCAACGGTGGCGCTCAACGCTGACTCGGTGTTCAGCTT 633
Db 430 CCAGTCCCAATGATGAACATCATCAACGGTGGCGCTCAACGCTGACTCGGTGTTCAGCTT 489
Qy 634 CAGGAATTCATGATCGCTCCCAATCGGTGCAAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
Db 490 CAGGAATTCATGATCGCTCCCAATCGGTGCAAGACCTTCTCTGAGGCTCTCCGCAACGGC 549
Qy 694 GCGAGGTCTACCAAGCACTGAAGTCCGTATCAAGGAAAGAGGCTGTCCACCGGACTT 753
Db 550 GCGAGGTCTACCAAGCACTGAAGTCCGTATCAAGGAAAGAGGCTGTCCACCGGACTT 609
Qy 754 GCGATCAGGCGGCTTCGCTCCCTTCGCTGGCTCCACCGTGGGCTCTTGACCTTATC 813
Db 610 GCGATCAGGCGGCTTCGCTCCCTTCGCTGGCTCCACCGTGGGCTCTTGACCTTATC 669
Qy 814 GTTGAGGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGCTCTGGAC 873
Db 670 GTTGAGGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGCTCTGGAC 729
Qy 874 GTTGCTTCTCTGAGTTCTTCAAGGAAGGCACTTCAACCCAGGCAAGGCTGGCAAGCTCC 933
Db 730 GTTGCTTCTCTGAGTTCTTCAAGGAAGGCACTTCAACCCAGGCAAGGCTGGCAAGCTCC 789
Qy 934 GCAGCTGAGATGGCAAGCTTACGCTGAGCTCGTTGACGGGTACCAATCGTCTCCATC 993
Db 790 GCAGCTGAGATGGCAAGCTTACGCTGAGCTCGTTGACGGGTACCAATCGTCTCCATC 849
Qy 994 GAGGACCACTGCAAGGAAGATGACTGGGAGGTTTACACCAACCTCACCGCAACCATCGC 1053
Db 850 GAGGACCACTGCAAGGAAGATGACTGGGAGGTTTACACCAACCTCACCGCAACCATCGC 909
Qy 1054 GACAAGTTTCAGATCGTTGGCAGCACTTCTTGCTCACCAACCTGAGCGCTGAAGGAG 1113
Db 910 GACAAGTTTCAGATCGTTGGCAGCACTTCTTGCTCACCAACCTGAGCGCTGAAGGAG 969
Qy 1114 GGCATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACGAGTCGGTACCCCTC 1173
Db 970 GGCATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACGAGTCGGTACCCCTC 1029
Qy 1174 ACCGAGACCTTCGACGCTGTGCATGCTCACCGCGCAGGCTACCTCCATGATCTCC 1233
Db 1030 ACCGAGACCTTCGACGCTGTGCATGCTCACCGCGCAGGCTACCTCCATGATCTCC 1089
Qy 1234 CACGCTTCGGGTGAGACCGGAGGACCAACCAATTCCTGACCTCGCAGTTGCACTCAACTGT 1293
Db 1090 CACGCTTCGGGTGAGACCGGAGGACCAACCAATTCCTGACCTCGCAGTTGCACTCAACTGT 1149
Qy 1294 GGCAGATCAAGACTGTGTCTCAGCAAGCTTCCGACCGTGTTCGCAAAAGTCAACAGCTT 1353
Db 1150 GGCAGATCAAGACTGTGTCTCAGCAAGCTTCCGACCGTGTTCGCAAAAGTCAACAGCTT 1209

Qy 1354 CTCCGCATCGAGCAGCTGCTTGGCGACGCCGCGCTCTACGAGGTGCGCAGCGCATTCCTCA 1413
Db 1210 CTCCGCATCGAGCAGCTGCTTGGCGACGCCGCGCTCTACGAGGTGCGCAGCGCATTCCTCA 1269
Qy 1414 CGCTTTCAGGGCTAA 1428
Db 1270 CGCTTTCAGGGCTAA 1284
RESULT 8
ACC45314
ID ACC45314 standard; DNA; 1278 BP.
XX ACC45314;
XX AC
XX AC
DT 17-JUN-2003 (first entry)
XX
XX Modified eno S330E encoding DNA SEQ ID NO:5.
XX
XX
KW FINE chemical; Coryneform bacteria; Escherichia coli; microorganism;
KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
KW pharmaceutical; food; animal feeding; eno; gene; ds.
XX
OS Corynebacterium glutamicum.
OS
XX
XX
XX
FT Key Location/Qualifiers
FT CDS 1..1278
FT /tag= a
FT /product= "modified eno S330E protein"
XX
XX W02003023016-A2.
XX
XX 20-MAR-2003.
XX
XX 11-SEP-2002; 2002WO-EP010174.
XX
XX 13-SEP-2001; 2001DE-01045043.
XX
XX (DEGS) DEGUSSA AG.
XX
XX Farwick M, Hermann T;
XX
XX WPI; 2003-354534/33.
XX P-PSDB; ABP97002.
XX
XX Microorganism useful for producing e.g. fine chemicals, has permanently
XX altered phosphorylatability protein, such that biosynthesis of fine
XX chemical synthesized by microorganism is increased compared to wild-type.
XX
XX Example 3; Page 83-84; 120pp; English.
XX
XX The present invention describes a microorganism (I), in which the
XX phosphorylatability of at least one protein has been permanently altered
XX such that the biosynthesis of at least one fine chemical synthesised by
XX the microorganism is increased compared to the wild type. Also described:
XX (1) use of a DNA (II) sequence coding for a protein which contains a
XX phosphorylation site, where the sequence contains such a mutation that
XX the protein is changed in its phosphorylatability for the production of
XX (1), or for the production of fine chemicals; and (2) a method for
XX producing fine chemicals or metabolites comprising using (I). (I) is
XX useful for producing fine chemicals or metabolites, such as amino acids,
XX vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
XX and vitamins produced using (I) can be used in human medicine, in the
XX pharmaceutical industry, food industry and in animal feeding. (I)
XX produces larger amount of desired fine chemical or a metabolite than the
XX wild type. The present sequence encodes a modified eno S330E from
XX Corynebacterium glutamicum, which is used in an example from the present
XX invention
XX
XX Sequence 1278 BP; 272 A; 397 C; 349 G; 260 T; 0 U; 0 Other;

Query Match	80.7%	Score 1273.2	DB 8	Length 1278
Best Local Similarity	99.8%	Pred. No. 1e-310		
Matches 1275	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Qy	151	GTGGCTGAATCATGCACGTAATTCGCTCGCGAAATTCCTCGATCTCCCGGGTAAACCAACC	210	
Db	1	GTGGCTGAATCATGCACGTAATTCGCTCGCGAAATTCCTCGATCTCCCGGGTAAACCAACC	60	
Qy	211	GTGAGGCAGAGGTTTTCTTGATGACGGTTTCCACGGTGTCCGAGGTGTTCATCCGGC	270	
Db	61	GTGAGGCAGAGGTTTTCTTGATGACGGTTTCCACGGTGTCCGAGGTGTTCATCCGGC	120	
Qy	271	GCATCCACCGGCTCCACGAGCTCATGAGCTCGGTGACGGTGGCGATCGCTACCTGGGC	330	
Db	121	GCATCCACCGGCTCCACGAGCTCATGAGCTCGGTGACGGTGGCGATCGCTACCTGGGC	180	
Qy	331	AAGGGCGTTTTGAAGGCAGTTGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGC	390	
Db	181	AAGGGCGTTTTGAAGGCAGTTGAAAACGTCAACGAAGAAATCGGCGACGAGCTCGCTGGC	240	
Qy	391	CTAGAGGCTGACATCAGCGCCTCATCGACGAGGAATGATCAAGCTGATCGCACCGC	450	
Db	241	CTAGAGGCTGACATCAGCGCCTCATCGACGAGGAATGATCAAGCTGATCGCACCGC	300	
Qy	451	AACAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGAAAGGT	510	
Db	301	AACAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGAAAGGT	360	
Qy	511	GCTGCTGATTCGACAGGCTTCCACTGTTCGGCTACATCGGTGGACCAACGACACGTT	570	
Db	361	GCTGCTGATTCGACAGGCTTCCACTGTTCGGCTACATCGGTGGACCAACGACACGTT	420	
Qy	571	CTTCCAGTTCCCAATGATGACATCATCAAGGTGGGCTCAGCTCAGCTCCGGTGTGAC	630	
Db	421	CTTCCAGTTCCCAATGATGACATCATCAAGGTGGGCTCAGCTCAGCTCCGGTGTGAC	480	
Qy	631	GTTCCAGGAATTCATGATCGTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	690	
Db	481	GTTCCAGGAATTCATGATCGTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	540	
Qy	691	GGCGCGAGGTCTACACGACGTGAAGTCGGTCAACAAGAAAGGCGCTGCCACCGA	750	
Db	541	GGCGCGAGGTCTACACGACGTGAAGTCGGTCAACAAGAAAGGCGCTGCCACCGA	600	
Qy	751	CTTGGCGATCAGGGCGGCTTCGCTCCTTCGGTGGGCTCCACCGTGAGGCTCTTGACCTT	810	
Db	601	CTTGGCGATCAGGGCGGCTTCGCTCCTTCGGTGGGCTCCACCGTGAGGCTCTTGACCTT	660	
Qy	811	ATCGTTGAGGCAATTCGAGAAAGGTGGCTTCAACCGGCAAGGACATCGCTCTTGCTCTG	870	
Db	661	ATCGTTGAGGCAATTCGAGAAAGGTGGCTTCAACCGGCAAGGACATCGCTCTTGCTCTG	720	
Qy	871	GAGGTTCCTTCCTGAGTTCTTTCAGGACGGCACCTACCACCTCGAAGGTGCCAGCAC	930	
Db	721	GAGGTTCCTTCCTGAGTTCTTTCAGGACGGCACCTACCACCTCGAAGGTGCCAGCAC	780	
Qy	931	TCGCGACGTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC	990	
Db	781	TCGCGACGTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC	840	
Qy	991	ATCAGAGACCCACTGCAGGAAGATGACTGGAGGGTTACACCAACCTCACCGCAACCATC	1050	
Db	841	ATCAGAGACCCACTGCAGGAAGATGACTGGAGGGTTACACCAACCTCACCGCAACCATC	900	
Qy	1051	GGCGACAAGGTTTCAGATCGTTGGCGACGACTTCTTTCGTTCCACCAACCTCAGCGCTGAAG	1110	
Db	901	GGCGACAAGGTTTCAGATCGTTGGCGACGACTTCTTTCGTTCCACCAACCTCAGCGCTGAAG	960	
Qy	1111	GAGGCGATCGTAAAGAGGCTGCCAACTCCATCTCTGGTTTAAGGTGAACAGATCGGTACC	1170	
Db	961	GAGGCGATCGTAAAGAGGCTGCCAAAGGATCTCTGGTTTAAGGTGAACAGATCGGTACC	1020	

Qy	1171	CTCACCAGACCTTCGACCGTGTTCGACATGGCTCACGGCAGGCTACACCTCCATGATG	1231
Db	1021	CTCACCAGACCTTCGACCGTGTTCGACATGGCTCACGGCAGGCTACACCTCCATGATG	1081
Qy	1231	TCCACCGTTCGGTGTGAGACCGAGGACACACACCATTCGCTGACCTCGCAGTTGCACCTCAAC	1291
Db	1081	TCCACCGTTCGGTGTGAGACCGAGGACACACACCATTCGCTGACCTCGCAGTTGCACCTCAAC	1141
Qy	1291	TGTGGCCAGATCAAGACTGGTGTCTCAGACGCTTCGACCGTTCGCAAGTTCGCAACGAC	1351
Db	1141	TGTGGCCAGATCAAGACTGGTGTCTCAGACGCTTCGACCGTTCGCAAGTTCGCAACGAC	1201
Qy	1351	CTTCTCCGCATCCAGCAGCTGCTTGGCGACGCGGGCTCTACGCAGTTCGACGGCATTC	1411
Db	1201	CTTCTCCGCATCCAGCAGCTGCTTGGCGACGCGGGCTCTACGCAGTTCGACGGCATTC	1261
Qy	1411	CCACGCTTTCAGGGCTAA	1428
Db	1261	CCACGCTTTCAGGGCTAA	1278
RESULT 9			
ID	ACC45313		
XX	ACC45313 standard; DNA; 1299 BP.		
AC	ACC45313;		
XX	17-JUN-2003 (first entry)		
DT	Modified endo DNA fragment S330E SEQ ID NO:4.		
DE	Genetic chemical; Corynebacterium bacteria; Escherichia coli; microorganism;		
XX	Genetically modified microorganism; metabolite; biosynthesis; amino acid;		
KW	vitamin; nucleoside; nucleotide; pigment; protein; human medicine;		
KW	pharmaceutical; food; animal feeding; endo; gene; ds.		
KW	Corynebacterium glutamicum.		
OS	Synthetic.		
XX	WO2003023016-A2.		
PN	20-MAR-2003.		
XX	11-SEP-2002; 2002WO-EP010174.		
PD	13-SEP-2001; 2001DE-01045043.		
XX	(DEGS) DEGUSSA AG.		
XX	Farwick M, Hermann T;		
PI	WPI; 2003-354534/33.		
XX	Microorganism useful for producing e.g. fine chemicals, has permanently		
XX	altered phosphorylatability protein, such that biosynthesis of fine		
XX	chemical synthesized by microorganism is increased compared to wild-type.		
PS	Claim 17; Page 82; 120pp; English.		
CC	The present invention describes a microorganism (I), in which the		
CC	phosphorylatability of at least one protein has been permanently altered		
CC	such that the biosynthesis of at least one fine chemical synthesised by		
CC	the microorganism is increased compared to the wild type. Also described:		
CC	(1) use of a DNA (II) sequence coding for a protein which contains a		
CC	phosphorylation site, where the sequence contains such a mutation that		
CC	the protein is changed in its phosphorylatability for the production of		
CC	(I), or for the production of fine chemicals; and (2) a method for		
CC	producing fine chemicals or metabolites comprising using (I). (I) is		
CC	useful for producing fine chemicals or metabolites, such as amino acids,		
CC	vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids,		
CC	and vitamins produced using (I) can be used in human medicine, in the		
CC	pharmaceutical industry, food industry and in animal feeding. (I)		
CC	produces larger amount of desired fine chemical or a metabolite than the		

CC wild type. The present sequence represents a modified eno fragment DNA
 CC sequence from Corynebacterium glutamicum, which is used in an example
 CC from the present invention
 XX
 SQ Sequence 1299 BP; 275 A; 407 C; 355 G; 262 T; 0 U; 0 Other;
 Query Match 80.5%; Score 1270.2; DB 8; Length 1299;
 Best Local Similarity 99.8%; Pred. No. 5.8e-310;
 Matches 1272; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 154 GCTCAATATGACGATTCCTCGCGAATTTCTCGACTCCCGGGTAAACCCACCGTC 213
 DB 10 GCTCAATATGACGATTCCTCGCGAATTTCTCGACTCCCGGGTAAACCCACCGTC 69
 QY 214 GAGCAGAGGTTTTCCTGGATGACGGTTCCACCGTGTCCAGTGTTCATCCGGGCA 273
 DB 70 GAGCAGAGGTTTTCCTGGATGACGGTTCCACCGTGTCCAGTGTTCATCCGGGCA 129
 QY 274 TCCACCGGCTCCACGAGGCTCATGAGCTGCGTGAACCGTGTCCAGTGTTCATCCGGGCA 333
 DB 130 TCCACCGGCTCCACGAGGCTCATGAGCTGCGTGAACCGTGTCCAGTGTTCATCCGGGCA 189
 QY 334 GGCCTTTTGAAGGCGAGTTGAACCGTCAACGAGAAATCGCGACGCTCGCTGGGCTA 393
 DB 190 GGCCTTTTGAAGGCGAGTTGAACCGTCAACGAGAAATCGCGACGCTCGCTGGGCTA 249
 QY 394 GAGCTGACGATCAGCGCTCATCGACGAAATGATCAAGCTTGTGGCAGCCGCAAC 453
 DB 250 GAGCTGACGATCAGCGCTCATCGACGAAATGATCAAGCTTGTGGCAGCCGCAAC 309
 QY 454 AAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGGTGTTCGAAAGGCTGCT 513
 DB 310 AAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGGTGTTCGAAAGGCTGCT 369
 QY 514 GCTGATTCGAGGCTCCCGACTGTTGCTGATCGTGTGACCAACGACGACGCTTCT 573
 DB 370 GCTGATTCGAGGCTCCCGACTGTTGCTGATCGTGTGACCAACGACGACGCTTCT 429
 QY 574 CCAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
 DB 430 CCAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
 QY 634 CAGGAATTCATGATCGCTCCAATCGGTGACGAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
 DB 490 CAGGAATTCATGATCGCTCCAATCGGTGACGAGACCTTCTCTGAGGCTCTCCGCAACGGC 549
 QY 694 GCGAGGTCTACACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
 DB 550 GCGAGGTCTACACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
 QY 754 GCGATGAGGCGGCTTCCGCTTCCGCTGAGTCCACCGTGTGAGGCTCTTGACCTTATC 813
 DB 610 GCGATGAGGCGGCTTCCGCTTCCGCTGAGTCCACCGTGTGAGGCTCTTGACCTTATC 669
 QY 814 GTTGAGGCAATTCGAGAGGCTGGCTTCAACCCGAGGAGGATCGCTTCTGCTCTGGAC 873
 DB 670 GTTGAGGCAATTCGAGAGGCTGGCTTCAACCCGAGGAGGATCGCTTCTGCTCTGGAC 729
 QY 874 GTTGCTTCTCTGATGTTCTTCAAGGAGGCGACCTTACCACTTCGAAAGTGGCCAGCACTCC 933
 DB 730 GTTGCTTCTCTGATGTTCTTCAAGGAGGCGACCTTACCACTTCGAAAGTGGCCAGCACTCC 789
 QY 934 GCAGCTGAGTGGCAAGGCTTACGCTGAGTCTGTTGACCGGTACCCCAATCGTCTCCATC 993
 DB 790 GCAGCTGAGTGGCAAGGCTTACGCTGAGTCTGTTGACCGGTACCCCAATCGTCTCCATC 849
 QY 994 GAGGACCCACTGACGAGGAGGATGATGAGGAGGTTTACACCACTTACCGCAACCATCGGC 1053
 DB 850 GAGGACCCACTGACGAGGAGGATGATGAGGAGGTTTACACCACTTACCGCAACCATCGGC 909
 QY 1054 GACAAGGTTGAGTGTGGCGACGATTTCTTCTGTCACCAACCTTGAGCGGCTGAAGGAG 1113
 DB 910 GACAAGGTTGAGTGTGGCGACGATTTCTTCTGTCACCAACCTTGAGCGGCTGAAGGAG 969

QY 1114 GGCATCGCTAAGAGGCTGCCAACTCCATCTCTGTTAAGGTGAACACGAGTCGGTACCCTC 1173
 DB 970 GGCATCGCTAAGAGGCTGCCAACTCCATCTCTGTTAAGGTGAACACGAGTCGGTACCCTC 1029
 QY 1174 ACCGAGACCTTTCGACGCTGTCGACATGCTCACCGCGAGGCTACACCTCCATGATGTC 1233
 DB 1030 ACCGAGACCTTTCGACGCTGTCGACATGCTCACCGCGAGGCTACACCTCCATGATGTC 1089
 QY 1234 CACCGTTCCGCTGAGAGCGAGGACACCACTTCTGACCTCGAGTTCGACTCAACTGT 1293
 DB 1090 CACCGTTCCGCTGAGAGCGAGGACACCACTTCTGACCTCGAGTTCGACTCAACTGT 1149
 QY 1294 GGCAGATCAAGACTGCTGCTCCAGCAGCTTCCACCGTGTCCGAAAGTACAACAGCTT 1353
 DB 1150 GGCAGATCAAGACTGCTGCTCCAGCAGCTTCCACCGTGTCCGAAAGTACAACAGCTT 1209
 QY 1354 CTCGCAATCGAGCAGCTGCTTGGCGAGCGCGGCTCTACGAGTTCGAGCGCATTCCTCA 1413
 DB 1210 CTCGCAATCGAGCAGCTGCTTGGCGAGCGCGGCTCTACGAGTTCGAGCGCATTCCTCA 1269
 QY 1414 CGCTTTCCAGGCTAA 1428
 DB 1270 CGCTTTCCAGGCTAA 1284
 RESULT 10
 ACA29854
 ID ACA29854 standard; DNA; 1275 BP.
 XX ACA29854;
 AC AC
 XX AC
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #11511.
 XX
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Corynebacterium diphtheriae.
 XX
 FN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU25984.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 17724; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
CC antiseptic nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antiseptic nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1275 BP; 268 A; 378 C; 354 G; 275 T; 0 U; 0 Other;

Query Match 57.6%; Score 908.6; DB 8; Length 1275;
Best Local Similarity 82.0%; Pred. No. 9.2e-219;
Matches 1046; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 151 GTGGCTGAATCATGACGATATTCGCTCGGAAATCTTCGACTCCCGGGTAACCCAAACC 210
DB 1 GTGGCTGACATATGACGATATTTGCTCGTGAATCTTGGACTCCCGGGTAACCCAAACC 60
QY 211 GTGAGGACAGAGTTTCTCGATGACGGTTCCACCGGTGTCGAGGTGTTCCATCCGGC 270
DB 61 GTTGAGCCGAAATTTCTTGGATGACGGATCCACCGCGTGTGGGGGTTCCTTCGGT 120
QY 271 GCATCCACCGGCTCCACGAGGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 330
DB 121 GCGTCCACCGGTGTTACGAGGCTCAGAGCTTCGCGACGGTGGCGAGGCTACCTCGG 180
QY 331 AAGGCGTTTGAAGGAGTGAAGAGTCAACGAGAAATCGCGAGGCTCGCTGGC 390
DB 181 AAGGCGTTTCAATGAGTGAAGAGTGAAGAGGAAATCGCTGACGCAATCGCTGGT 240
QY 391 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGCAGCGCC 450
DB 241 GCAGAGCCGACGATCAGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTCGGTGTTTCCATGGCTGTTGCAAGGCT 510
DB 301 AACAAAGTCTGCTCGCGGCAACGCTATCTTGGGTGTTATCCATCGCGGTAGCTAAGGCT 360
QY 511 GCTGCTGATCCGAGGCTCCACTGTTCCGCTACATCGGTGACCAAGCAACGACAGTT 570
DB 361 GCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 571 CTTCAGGTTCAATGATGAATCATCAACGGTGGGCTCAGCGTGAATCGGTGTTGAC 630
DB 421 CTTCAGGTTCTATGATGAATGTTAAACGGTGGGCAACGCTGACTCGGGGTTGAT 480
QY 631 GTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
DB 481 GTTCAGGAGTTCATGATGCTCCTATCGGTGCGAGTCTTCTCTGAGGCTCTGCGCATG 540
QY 691 GGCGGGAGGCTTACACGCACTGAAGTCCGTCATCAAGGAAAGGGGCTGTCACCGGA 750
DB 541 GGTGACAGGCTTACCACTCTTGAAGTCCGTTAAGTCAAGGGGACTTTCACCGGC 600
QY 751 CTTGGCGATGAGGCGGCTTCGCTCTTCCGCTCGGCTCCACCGCTGAGGCTCTTGACCTT 810

DB 601 CTGCGGACGAAGGTGGTTTCGCACCTTCTGTTGAGTCCACCAAGGAGCTCTCGACCTC 660
QY 811 ATCGTTGAGCAATCGAGAGGTGGCTTACCCTGAGGCAAGGACATCGTCTTGTCTGTG 870
DB 661 ATCGTTGAGCAATCGAGAGGTGGCTTACCCTGAGGCAAGGAGTCTGACATCGCACTTC 720
QY 871 GAGTGTCTTCCCTGAGTTCTTCAAGGACGCGACCTACCACTTCAAGGTTGGCGCACAC 930
DB 721 GAGTGTCTTCCCTGAGTTCTTCAAGGATGGCAAGTACCACTTCAAGGTTGGCGCACAC 780
QY 931 TCCGCAAGCTGAGATGGCAAGCTTACGCTGAGCTTCCGCTGAGCTACCACTTCAAGGTT 990
DB 781 ACCGCTGAGGAGATGGCAAGCTTACGCTGAGCTTCCGCTGAGCTACCACTTCAAGGTT 840
QY 991 ATCGAGACCCATCGCAGGAGATGATGAGGAGGTTTACCACTTCAAGGTTGGCGCACAC 1050
DB 841 ATCGAGACCCATCGCAGGAGATGATGAGGAGGTTTACCACTTCAAGGTTGGCGCACAC 900
QY 1051 GCGCAAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1110
DB 901 GGTGACAAAGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTT 960
QY 1111 GAGGAGATCGCTAAGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1170
DB 961 GAAGGATCGAGAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1020
QY 1171 CTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTT 1230
DB 1021 CTGACCGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTT 1080
QY 1231 TCCACCGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1290
DB 1081 TCCACCGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1140
QY 1291 TGTGGGCAAGTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1350
DB 1141 TGTGGGCAAGTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1200
QY 1351 CTTCTCCGATCGAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1410
DB 1201 CTTCTCCGATCGAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAGGAGGTTTCAAG 1260
QY 1411 CCAGGCTTTTCAGGCG 1425
DB 1261 CCAGGCTTTTCAGGCG 1275

RESULT 11

AD56353
ID AD56353 standard; cDNA; 1269 BP.

XX AC AD56353;

XX AC AD56353;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #8340.

XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX Prokaryotic essential gene #19736.
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Mycobacterium avium.
OS
XX W0200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR P-PSDB; ABU34209.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 25949; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid;
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: the sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

Sequence 1287 BP; 214 A; 424 C; 475 G; 174 T; 0 U; 0 Other;
Query Match 41.0%; Score 647; DB 8; Length 1287;
Best Local Similarity 69.9%; Pred. No. 8.8e-153;
Matches 888; Conservative 0; Mismatches 380; Indels 3; Gaps 1;
151 GTGGCTGAATCATGCACGTATTCGTCGCGAAATTCGACTCCCGCGGTAACCCAC 210

1 GTCCCATATTCGAGCAGGTTCGGGCCCCGGAGATCCTCGACTCCCGCGGTAAACCGACA 60
211 GTGAGCGCAGAGGTTTTCCTGGATGAGGTTCCACAGGTGTCCAGAGTGTTCATCCGGC 270
61 GTGAGGTGAGATCGCCCTTCACCGAGGAACATTCGCCCGCGCGGGTTCGCGTCCGG 120
271 GCATCCACCGCGCTCCACGAGGCTCATGAGTCGGTGAGGTGCGGTGATCCCTACCTGGGC 330
121 GCCTGACCGGTGAGCAGAGGCTGTGGAATCTCGCGACGCGGGGAGCGGTACGGCGGC 180
331 AAGGCGTTTGAAGCGAGTTGAAAACGTCACAGAAAGAAATCGCGCGACGAGCTCGCTGGC 390
181 AAGGCGTGCAGAAAGCGGTGCAAGCGGTGCTGACGAGATCGCCCGCGGTGATCGGG 240
391 CTAGAGGCTGAGATCAGAGCGCTTCATCGAGGAAGCAATGATCAAGCTTGATGGCAGCGCC 450
241 CTGAACGCCACGACACGAGCGCTGGTGGACAGCGCTCGTGACCTGGACCGGACCGCCG 300
451 AACAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGCTGCTGTTCGAAAGGCT 510
301 GACAAGTCGAGGCTGGCGGCAACGGATCCTGGGCTGTGCTGGCGGTGGCGGCAAGGCG 360
511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGCTACATCGGTGAGCAACGACACAGTT 570
361 GCCCGGATTCGCGGAGTTGCCGCTGTTCCGCTACTCTGGCGGGGCCCAACGCGCACATC 420
571 CTTCCAGTTCATGATGAAATCATCAACGGTGGGCTCACGCTGACTTCGCTGTTGAC 630
421 CTGCGGTGCGATGATGAACATCCTCAACGGCGGTGCGCACACCGCGCTCGAC 480
631 GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
481 ATCAGAGATTCATGTCGCGCCGATCGCGCGGAGTTTCGCGGAGGATTCGCGTGG 540
691 GCGCGGAGGTCTACCAACGACATGAAGTCCGTTCATCAAGAAAGGCGCTGTCCACCGA 750
541 GGTGCGAGGTGTATCACTCGTGAAGTCCGTGCTGAAGAAAGGAGGCGCTGAGCACCGGC 600
751 CTTGGCGATCAGGCGGCTTCGCTCTTCGCTGGCTCCACCGTGGAGGCTCTTGACCTT 810
601 CTGGGCGACGAGGCGGCTTCGCGCGCGAGCTGGCGGCGCACACCGCGGCGCTGGACCTG 660
811 ATCGTTGAGCAATTCGAGAAAGGCTGGCTTCACCCAGGCAAGACATCGCTCTTGTCTG 870
661 ATCGCGCGGCGCATCGAATCGCGCGGCTTCAAACTCGGACCGAGCTGGCGCTGGCCCTC 720
871 GAGCTTGCTTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCGAAAGTGGCCAG 927
721 GACGCGCGCGCCACCGAGTTCTACAGCGACGGCACCGGCTACAAGTTTCGAGGGGAGCACC 780
928 CACTCCGACGTGAGATGGCAACAGTTTACGCTGAGCTCGTTGACGGGTACCCCAATCGTC 987
781 CGCACCCCGGACGAGATGGCGGAGTTCTACGCGCGGCTGCTCGCGCGGTATTCGCTGGTG 840
988 TCCATCCAGGACCCACATCGCAGGAAGATGACTGGGAGGTTTACACCAACCTCACCGCAACC 1047
841 TCCATCGAAGACCCGCTGTCGAGATGATTGGAGCGGCTGGGCGGCGCTGACCGCGTCG 900
1048 ATCGCGCAAGGTTTCAGATCGTTGGCGAGCATTTCTTCGTACCAACCCCTGAGCGGCTG 1107
901 ATCGCGCACCGGTGACGCTCGTGGCGGACGACGCTTCGTCGACGAAACCCCGAAGCGCTG 960
1108 AAGGAGGGCATCGCTAAGAGGCTGCCAATCCCATCTGTTTAAAGGTGAACCATCGGT 1167
961 GAAGAGGGTATCGGAGAAAGGGCGTCGAAATTCGTTGCTGGTCAAGGTGAATCATCATGGC 1020
1168 ACCCTCACCGAGACCTTCGACGCTGTGAGATGGCTCACCGCGAGGCTACACCTCCATG 1227
1021 ACGTGCACCGAGACGCTGGAGCGCGTTCGCGTGGCCCAACACACAGCGGCTACCGCACATG 1080
1228 ATGTCACACCGTTCGGGTGAGACCGGAGACACCACTTCTGACCTCGCAGTTGCACTC 1287

Db 1081 ATGAGCCACCGCAGCGGTGAACACGAGACACACGATCGCCGACCTGGCGGTGGCCCTC 1140
 Qy 1288 AACCTGTGGCGAGATCAAGACTGTGCTCCAGCAGCTTCGAGCCGTGTGCAAACTACAAC 1347
 Db 1141 GGACAGCGGCGAGATCAAGACCGGAGCGCGCGCCGCGAGCGGCTGGCCNAGTACAAC 1200
 Qy 1348 CAGCTTCTCCGCGATCGAGCAGCTGCTGTGGCGACCGCGCGGTCTACGAGGTCGCGAGCGCA 1407
 Db 1201 CAGCTGTCTGGGATCGAGGAGCGCTCGCGCAGCGCGCGCTTACGCGCGCGACCTGGCC 1260
 Qy 1408 TTCCACAGCTT 1418
 Db 1261 TTCCCGCGGTT 1271

RESULT 13

AAS59564/C
 ID AAS59564 standard; DNA; 30159 BP.

AC AAS59564;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein encoding DNA #59.

DE SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

OS WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris.

XX Claim 1; SEQ ID NO 59; 1069pp; English.

XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
 CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by P. acnes. The disorders include
 CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
 CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
 CC infections of bone, joints and the central nervous system, however it is
 CC particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of P. acnes in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for P. acnes proteins. These antibodies can be
 CC used to downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU54183-AAU54498 and AAU67555-AAU67557. Note: The

CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 30159 BP; 6067 A; 9304 C; 8907 G; 5878 T; 0 U; 3 Other;
 Query Match 40.7%; Score 642.4; DB 4; Length 30159;
 Best Local Similarity 69.3%; Pred. No. 3.6e-151;
 Matches 890; Conservative 0; Mismatches 391; Indels 3; Gaps 1;
 Qy 152 TGGCTGAATCATGCACGTATTTCGCTCGCAAAATTCCTGCAGTCCGCGGTAAACCAACCG 211
 Db 1539 TGGCAACCATCGAATTCATCGAAGCCGTCGAGATCCTCGATTTCCCGGGCAACCCGACCG 1480
 Qy 212 TCGAGGCGAGAGGTTTTCTCGATGACGTTCCCAACGGTTCGACAGGTGTTCCATCCGGCG 271
 Db 1479 TTGAGGTCGAGATGATCCTCGACGACGCCACCCAGGCCGCGCTGCGGTTCTTCGGGCG 1420
 Qy 272 CATCCACCGCGCTCACGAGGCTCATGAGTGCCTGACGG--TGGCGATCGCTACCTGG 328
 Db 1419 CTTTCGACCGGTTCAGTTCGAGGCCGTTGAGCTGCTGACGGCGATAAGAAGGTTACTCCG 1360
 Qy 329 GCAAGGCGTTTTGAAGCGAGTTGAAAACGTCACGAAAGAAATCGGCGACGAGTCCGTG 388
 Db 1359 GTAAGGGTGTCTCAAGGCTGTTGAGAAACGTCACGAGAAAGATCGCCGAGAGGTGCTCG 1300
 Qy 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAAAGCAATGATCAAGCTTCATGCGACCG 448
 Db 1299 GCTGTGACGAGTGACGCGCATATCGACAGATCATGATCGAGCTCGACGATCCG 1240
 Qy 449 CCAACAAAGTCCCGCTGGGTGCAAAACGCAATCTTGGTGTTCATGCTGCTTGCACAAAG 508
 Db 1239 ACAACAAAGGCAAGTTGGTGTCTAAACGCCATCTTGGTGTCTCCCTGGCTGCTCATG 1180
 Qy 509 CTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGAGCAACCAACGACGACG 568
 Db 1179 CTGCTGCTGACTGTGCGGAGCTCCCGCTGTACCAAGTACCTCGGCGGACCGAACTCCACG 1120
 Qy 569 TTCTTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACCTCCGCTGTTG 628
 Db 1119 TGTGCCCCGTTCCCAATGATGAACATCTCTCAACGGTGGTCCCGCCAGCGGATTTCCGACGTTG 1060
 Qy 629 ACCTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGAGACCTTCTCTGAGGCTCTCCGCA 688
 Db 1059 ACATCCAGGATTCATGATTTGCCCGGATCGGCGGGAATCCTTTAAGCAGGCTCTACGAGT 1000
 Qy 689 ACGGCGGAGGTTCTACCGCATGAAAGTCCGTCATCAAGGAAAGGCGCTGTCACCG 748
 Db 999 GGGCGCTGCGGTTTACCACTCCCTCAAGAAGGTTCTCAAGGCAAGGGCTTGGCTACCG 940
 Qy 749 GACTTGGCGATGAGGCGGCTTCGCTCTTCCGTCGGCTCCACCGTGAGGCTCTTGACC 808
 Db 939 GTCTGGGTGACGAGGCGGCTTTCGCCCCCAACCTTCCACAGCAACGCGCGCTTTGGACC 880
 Qy 809 TTATCTGTTAGGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGGAGATCGCTCTTGCTC 868
 Db 879 TCATCTTACGCCCATCAAGCCGCGGCTTCGAGCGGGTAAAGACGTCGCGCTCGGCC 820
 Qy 869 TGGAGGTTGCTTCTCTGAGTTCTTCAAGGACGCACTACCACTTCGAAGGTGGCCAGC 928
 Db 819 TTGATGTGCTGCTTCCGAATTCCTTCGAGGACGGCAAGTACACGTTTCGAGGGCCAGGCTA 760
 Qy 929 ACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGCGTACCAATCGTCT 988
 Db 759 AGACCTCGGCTGAGATGATCGCGTACTACGAAGGCCCTCATCGCCAAGTATCCGCTGGTCT 700
 Qy 989 CCATCGAGGACCCACTCGAAGATGATCACTGGGAGGTTTACACCAACTCACCCGCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGACGAGGAGGATTTGGGACGGTTGGGCCGAGTTTACGAGAGAC 640
 Qy 1049 TCGGCGCAAGGTTTTCAGATCGTTGGCGACGACTTCTTCGTACCAACCCCTGAGCGCTTGA 1108
 Db 639 TTGCTGAGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCCTCACTAAACCCGAGCGCTTG 580

Db 819 TTGATGTGCTGCTTCCGAATTTCTCGAGGACGGCAAGTACACGCTTCGAGGGCCAGGCTA 760
 Qy 929 ACTCCGAGCTGAGATCGCAACGCTTTACGCTGAGCTGCTTACGCGGTACCAATCTGCT 988
 Db 759 AGACTCGGCTGAGATATCGGTAATCTACGAGGCTTATCGCAAGTATCCGCTGCT 700
 Qy 989 CCATCGAGGACCCACTCGAGGAAGATGACTGGAGGGTTACCAACCTCACCCCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGGACGAGGAGATTGGGACGGTTGGCCGAGTTACAGGAAGC 640
 Qy 1049 TCGCGACACAGGTTACAGATCGTTGGGACGACTTCTTCTGTCACCAACCTGAGCGCTGA 1108
 Db 639 TTGTTGAGAGATCCAGATCGTTGGGACGACTTGTTCCTCACTAACCCGAGCGCTTG 580
 Qy 1109 AGAGGGCATCGCTAACGAAGCTGCCAATCTCATCTCTGGTTAAGGTGAACAGATCGGTA 1168
 Db 579 CCAAGGGAATCGAGACCAAGGCTGCCAAGCCCTGCTCTCAGGTGACAGATTGGCT 520
 Qy 1169 CCCTCAACGAGACTTCGACGCTTCGACATGCTGCTACCGCGAGGCTTACACTTCCATGA 1228
 Db 519 CCCTCTCGGAGACTATCGACGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 460
 Qy 1229 TGTCCACCGTTCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1288
 Db 459 TGTCCACCGTTCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 400
 Qy 1289 ACTGTGGCCAGATCAAGACTGCTGCTCCAGCAGCTTCGACCGGTTCGCAAGTACAAAC 1348
 Db 399 CGACTGTCAGATTAAATCCGCTGCCCGCCCGGCGGAGCGCATGCCAAGTACAAAC 340
 Qy 1349 AGCTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTTACGCGAGTTCGAGCGCAT 1408
 Db 339 AGTTGCTGCTGATTGAGGAGGAACTGGGCGAGTTCGCGGAGTACGCGGCTTCGCGCT 280
 Qy 1409 TCCACGCTTTCAGGGCTAAATAA 1432
 Db 279 TCCGCGCTTCCAGGCTGACTGA 256

RESULT 15

ABX09144
 ID ABX09144 standard; DNA; 45190 BP.

XX AC ABX09144;

XX DT 08-APR-2003 (first entry)

XX DE Mycobacterium tuberculosis H37Rv BAC clone BAC-Rv267.

XX KW Mycobacteriosis; survival; virulence; protective antigen; vaccine;
 mycobacterial disease; tuberculosis; leprosy; ds; cosmid.

XX OS Mycobacterium tuberculosis.

XX PN W0200274903-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002MO-IB001973.

XX PR 22-FEB-2001; 2001US-0270123P.

XX PA (INSP) INST PASTEUR.

XX PI Cole S;

XX DR WPI; 2002-759885/82.

XX PT Identifying and selecting genes for survival or virulence of mycobacteria
 by a comparative genomic analysis of the sequences of Mycobacterium
 tuberculosis and M. lepreae.

XX PS Disclosure; Fig 8; 874pp; English.

XX

CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC lepreae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a Mycobacterial cosmid DNA
 CC sequence used in the method of the invention

SQ Sequence 45190 BP; 7490 A; 14522 C; 15204 G; 7974 T; 0 U; 0 Other;

Query Match 40.1%; Score 632.6; DB 6; Length 45190;
 Best Local Similarity 68.4%; Pred. No. 1.2e-148;
 Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGCTGAATCATGACAGTATTCGCTCGGAAATTCGACTCCCGCG 199
 Db 19931 AGGAGAACCCAGTGCCTGATTATCGAGCAGGTTAGGGCCCGAGAGATCCTCGATTCCCGCG 19990
 Qy 200 GTAACCCCAACCGTCGAGGCGAGAGGTTTCTCTGATGACGGTTCCACGGTGTCCGAGGTG 259
 Db 19991 GCAACCCGACGGTGGAGTTCGAGTGGCGCTTATCGAGCGGACATTCGCCCGGCGCGG 20050
 Qy 260 TTCCATCCGCGCATCCACCGCGTCCACGAGGCTCATGAGTCCGTGACGGTGGCGATC 319
 Db 20051 TGGCGTCCGCGCTCCACCGGCGAGCAGAGGCGTTCGAGTTCGCGGACGCGCGCGATC 20110
 Qy 320 GCTACTGGCGCAAGGGCGTTTGAAGGCGATTGAAACGTCACAGAGAAATCCGCGACG 379
 Db 20111 GCTACGGCGCAAGGGCGTTCGAAAGCGCTGCGAGGCTGTTCTTGATGAGATCGGCGCG 20170
 Qy 380 AGCTCGCTGCGCTAGAGGCTGACGATCAGCGCCCTCATCGAGAGCAATGATCAAGCTGTG 439
 Db 20171 CCGTCATCGACTCAACCGCGAGCAGCCGATGTTGTCGACGAGCGCTGTTGACCTAG 20230
 Qy 440 ATGGCACCGCAACAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATGGGTG 499
 Db 20231 ACGCACCCCGCAAGTCCCGCTGGGCGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 20290
 Qy 500 TTGCAAGGCTGCTGCTGATTCCGCGAGCTCCCACTGTTCCGCTACATCCGTTGGACCA 559
 Db 20291 TTGCCAAGCGCGCGCGGATTCGCGGAGCTGCCGTTGTTCCGTTATGTTCGGGGGCGCA 20350
 Qy 560 ACGCACAGCTTCTTCAGTGTCCAAATGATGAACATCATCAACGTTGGGCTCACGCTGACT 619
 Db 20351 ACGGCGACATTCGCGCTACCGATGATGAACATCTTCACGCGGCGGACACCGCGATA 20410
 Qy 620 CCGGTGTGACGTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGAGACTTCTCTGAGG 679
 Db 20411 CCGGTGTGACATTCAGAGATTTCATGTTGGCGCAATTGGCGCGCCAGCTTCGTCGAGG 20470
 Qy 680 CTCTCCGCAACGGCGCGAGGTCTACACGCACTGAAGTCCGTATCAAGAGAAAGGGCC 739
 Db 20471 CGTTGCGCTGGGCGCTGAGGTGTACACGCGCTCAAGTCGGTCTGAAAAAGAGGGGC 20530
 Qy 740 TGTCCACCGGACTTGGCGGATGAGGGCGCTTCGCTCTCTTCGTCGGTCCACCGTGGAGG 799
 Db 20531 TGTCCACCGGCTGGGCGACGAAGCGGCTTCGCCCCGGGATGTGCCGGGACACCGCGG 20590
 Qy 800 CTCTTGACCTTATCTGTTGAGGCAATCGAGAGGCTGGCTTACCCCCAGGCAAGACATCG 859
 Db 20591 CGTTGACCTGATCAGCGGCGCCATCGAGTCGGCGGCTTCGCGACCCGCGCGCGAGCTG 20650
 Qy 860 CTCTTGCTTGAGGCTTGTCTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG 916
 Db 20651 CGTTGGCCCTGGACGCGCGCGCCACCGAGTTCTTTCACGAGCGGACCGGCTTCTTCG 20710

Qy	917	AGGTGCCAGCACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGCGT	976
Db	20711	AGGGCACACCCGTACCGCAGACGAGATGACCGAGTTCTACGGGGGCTCTCGGCGCT	20770
Qy	977	ACCCAATCGTCTCCATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTACACCAACC	1036
Db	20771	ACCGCTGGTGTGATCGAAGACCCACTGTCCGAGAGACGATTGGGACGGCTGGGCGCGC	20830
Qy	1037	TCACCGCAACCATCGGCGACAAGGTTGAGATCGTTGGCGAGGACTTCTCGTCACCAACC	1096
Db	20831	TGACGGCTCGATCGGTGACCGGTGCAATCGTCGGCGACGACATCTTTGTCAACCAATC	20890
Qy	1097	CTGAGCGCTGAAGGAGGGCATCGCTAAGAGGCTGCCAACTCCATCCTGTTAAGTGA	1156
Db	20891	CCGAGCGGCTCGAGGAGGGCATCGAACGGGCGTGGCAATGCGTTGCTGTCAGGTGA	20950
Qy	1157	ACCAGATCGGTACCCCTCACCGAGACCTTTCGACGCTGTCGACATGGCTCACCGCGCAGGCT	1216
Db	20951	ACCAGATCGGACGTTGACCGAGACACTCGACGCGGTACGCTGGCTCACCGCGCGAT	21010
Qy	1217	ACACCTCATGATGCCACCGTTCCGGTGAGACCGAGGACACCACTTGTGACCTCG	1276
Db	21011	ACCGACGATGATCAGTCAACCGCAGTGGCGAGACGAGGACACCAATGATCGCCACCTCG	21070
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCACGTTCCGACCGTCTCG	1336
Db	21071	CGGTGGCCATCGGCAGCGGGCAGATCAAGACGGGCGGCTGCTCGCAGTGAGCGCGT	21130
Qy	1337	CAAAGTACAACCAAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGACCGCGGCTCTACGCAG	1396
Db	21131	CAAAATACAACCAAGCTGCTCGGATCGAAGAGGCGCTTGGCGACCGCGCGCTACCGG	21190
Qy	1397	GTCGAGCGGATTTCCACGCTTTGAGGCTTAAATAAAGCGCT	1439
Db	21191	GCGACCTGGCATTTCTCGTTCGCTGCGGAGACGAAATAGGT	21233

Search completed: September 28, 2005, 17:01:31
Job time : 801.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:03:14 ; Search time 4756.5 Seconds
(without alignments)
12628.070 Million cell updates/sec

Title: US-10-728-947-1

Perfect score: 1578

Sequence: 1 ggctgggatatgggtagt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsl1.*
9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	407.6	25.8	1522	4	BM320865
2	397.6	25.2	2491	8	BH770842
3	386.8	24.5	1626	3	CR685174
4	385.2	24.3	1643	3	CR728296
5	384.2	24.3	1505	4	BM321196
6	382	24.2	1623	3	CR679338
7	377.4	23.9	1613	3	CR698859
8	377.4	23.9	1615	3	CR650019
9	377	23.9	1637	3	CR728153
10	369.6	23.4	1551	3	CR685932
11	366.8	23.2	1641	3	CR672414
12	364.6	23.1	1582	3	CR656482
13	341.8	21.7	2021	8	BH770828
14	328	20.8	1551	3	CR658829
15	327	20.7	1407	3	CR733357
16	326.4	20.7	1476	3	CR726536
17	326.4	20.7	1563	3	CR657796
18	326.4	20.7	1587	3	CR675445
19	326.4	20.7	1592	3	CR664987
20	326.4	20.7	1592	3	CR731975
21	325.6	20.6	1588	3	CR731270
22	325.2	20.6	1566	3	CR729213
23	324.8	20.6	1518	3	CR675234
24	324.8	20.6	1536	3	CR670482

ALIGNMENTS

RESULT 1

BM320865

LOCUS

DEFINITION

rockefeller.0.878 Mastigamoeba balamuthi lambda ZAP II library

Mastigamoeba balamuthi cDNA similar to enolase (EC 4.2.1.1), mRNA

sequence.

ACCESSION

BM320865

VERSION

EST.

KEYWORDS

Mastigamoeba balamuthi

SOURCE

Mastigamoeba balamuthi

ORGANISM

Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

REFERENCE

1 (bases 1 to 1522)

AUTHORS

Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

Gordon,P., Durufle,L., Gaasterland,P., Lopez,P., Muller,M. and

Philippe,H.

The analysis of 100 genes supports the grouping of three highly

divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 1522 Std Error: 0.00

POLYA=Yes.

FEATURES

source

1..1522

/organism="Mastigamoeba balamuthi"

/mol_type="mRNA"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match

Best Local Similarity

Matches 787; Conservative

25.8%; Score 407.6; DB 4; Length 1522;

Pred. No. 1.3e-98;

Mismatches 464; Indels 51; Gaps 5;

Oy 143 AGCCACAGTGGTGAATCATGCACGTATTCGTCGGAATTCGACTCCCGCGTA 202

Db 30 AAGCAACAATGTGACCATCAAGAGCTCTTCGCCCGGAGACTCTCGTGCA 89

Oy 203 ACCCAACCGTCGAGGAGAGGTTTTCTCTGGATGACGGTCCACGGGTGTCGACGGTGTTC 262

CR669496 Tetraodon
CR659851 Tetraodon
CR678580 Tetraodon
CR676296 Tetraodon
CR658992 Tetraodon
CR650486 Tetraodon
CR685781 Tetraodon
CR664375 Tetraodon
CR676706 Tetraodon
CR6724795 Tetraodon
CR673577 Tetraodon
CR729481 Tetraodon
CR723707 Tetraodon
CR726056 Tetraodon
CR728206 Tetraodon
CR646460 Tetraodon
CR653599 Tetraodon
CR649337 Tetraodon
CR656841 Tetraodon
CR678734 Tetraodon
CR723775 Tetraodon

1522 bp mRNA linear EST 03-JAN-2002
rockefeller.0.878 Mastigamoeba balamuthi lambda ZAP II library
Mastigamoeba balamuthi cDNA similar to enolase (EC 4.2.1.1), mRNA

sequence.

ACCESSION

BM320865

VERSION

EST.

KEYWORDS

Mastigamoeba balamuthi

SOURCE

Mastigamoeba balamuthi

ORGANISM

Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

REFERENCE

1 (bases 1 to 1522)

AUTHORS

Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

Gordon,P., Durufle,L., Gaasterland,P., Lopez,P., Muller,M. and

Philippe,H.

The analysis of 100 genes supports the grouping of three highly

divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 1522 Std Error: 0.00

POLYA=Yes.

FEATURES

source

1..1522

/organism="Mastigamoeba balamuthi"

/mol_type="mRNA"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

```
Db 90 ACCACCGCTCGAGGTTGACCTTACACACGAGAGGCGCTCTTTCCG---CTCGCGCGTCC 146
Qy 263 CATCCGCGCATCACCGCGGTCCACGAGGCTCATGAGCTGCGTGAAGCGTGGCGA---TC 319
Db 147 CTTCTGTGCTCGACTGGGATCTATGAGGCTTCGAGCTCCGTCGCGGCGACAAGTCCG 206
Qy 320 GCTACTCTGGCAAGGGCGTTTGAAGGCACTTGAAGAACTGAAACGTCACGAGAAATCGGCGAG 379
Db 207 GCTACTCTGGCAAGGGGTGCTGAAGGCGCTCGAGAACGTCACAAAGATCTCGCGCCCA 266
Qy 380 AGCTCGCTGAGGCTGAGGCTGACGATCAGCGCTCTACGAGCAAGCAATCATCAAGCTTG 439
Db 267 AGCTGATCGGCTCGAGCTCACCAAGCAGGCGGAGATCGACAGGCTCATGCTCOAGATCG 326
Qy 440 ATGACACGCAACAAGTCCCGCTCGGCTGAAACGCAATCCTTGGTGTTCATCGGCTG 499
Db 327 ACGGCACTGAGAACAGACCACTCGGCGGCAAGGCAATCCTCGGCTGCTCGATGTGG 386
Qy 500 TTGCAAAAGGCTGTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGT----- 552
Db 387 TCTCGCTGCGGCGCGCGTTCAGGGCGCTGCGCCCTCTACCGCTACATCGCGAGCTGT 446
Qy 553 --GACCAACGCAACAGTCTTCCAGTTCATGATGAACATCATCAAGGTTGGCGCTC 610
Db 447 CGGCAACAAGTCCGCTCATGCTCTGCGCTGCTTCAAGCTCATCAAGCGGCGTGAAG 506
Qy 611 ACGTGTACTCGGCTGTTGAGCTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGACCT 670
Db 507 ACGGGGCAACAAGCTCGCCATCGAGGATTCATGATCGCCCAAGCGGCGGCGACCTGT 566
Qy 671 TCTCTGAGGCTCTCGCAACGCGCGGAGGCTTACAGCGCACTGAAGTCCGTCTCATGAGG 730
Db 567 TCCACGAGGCGCTCGCATGCGCGCGGAGAGTACCAACCTGAAGCTGTGTCATCAAGA 626
Qy 731 AABAGGGC-----CTGTCCACCGGACTTGGCGATGAGGGCGGTTCGCTCCTTCCGTCG 784
Db 627 AGAGGTACGGATGAGCGGACCAAGTCGCGGACGAGGGTGGTTCGCCCCCAAGATCC 686
Qy 785 GCTCCACCGGTGAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGGCTGGCTTACCC 844
Db 687 AGGCCAACCAAGAGGCGCTCGAGCTCATCGTTGAGGCCATCAAGCAGCGCGGTACACTG 746
Qy 845 CAGGCAAGGACATCGCTCTGCTGAGCTGCTGCTCTGAGTTCTTCAAGGACGGCA 904
Db 747 GCAAGATCGAGATCGGATGAGCGTTGCGCGCTCGTGTCTCGGAGCGCCAAAGAGTGA 806
Qy 905 CTTACCACTTCGAMGTTGGCCAGCACTCCGCGACTGAGA----- 943
Db 807 AGTAGCACTCGGCTTCAAGTTCGCGCGGACAGAGACCCCGACATGCTGCTCTCGG 866
Qy 944 -----TGGCAACGTTTACGCTGAGCTGCTGTTGACGCGTACCCAACTGCTCCATCG 994
Db 867 GCGAGGCGCTGATCAAGCTTACGAGGAGTGGACCTCGAAGTACCGGATCTGCTGCTG 926
Qy 995 AGGACCACTCGAGGAAGATGACTGGAGGGTTACACCAACCTCACCGCAACCTACCGG 1054
Db 927 AGGACCCCTTCGACAGGACCACTGGGCCA CTTACACCAAGTTTACCGAGCTGATCGGCA 986
Qy 1055 ACAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCAACCAACCTGAGCGCTGAAGGAGG 1114
Db 987 ACAGATCCAGATCTCGCGGACGACTCTGCTGTCACCAACCCCAAGCGCATCGTCGAGG 1046
Qy 1115 GCATCGCTAAGAGGCTCGCAACTCCATCTCGTGTTAAGGTGAACAGATCGGTACCTCA 1174
Db 1047 CCGCAACAAGAGGCGCTGCAAGCGGCTGCTCAAGCTCAACAGATCGGCTCGGTCA 1106
Qy 1175 CCGAGACCTTCGACGCTGTCGATGCTCACCGCGAGGCTTACACCTCCATGATGTCC 1234
Db 1107 GCGAGGCGTTCGAGGCGTCCGCTGCGCGGAGGTCAACTGGGGCGTCACTGCTGTCG 1166
Qy 1235 ACCGTTCCGGTGGAGCCGAGGACACCACTTGTGACCTCGCACTTGCATCAACTGTG 1294
```

```
Db 1167 ACCGTCGGCGAGACTGAGGACGGCTTCATCGCGGACCTCGTCTGGTCTGGGCTGG 1226
Qy 1295 GCAGATCAAGACTGGTGTCTCAGCAGCTTCCGACCGTGTCCGAAAGTACACCGAGCTTC 1354
Db 1227 GCAGATCAAGACTGGTGTCCCTCGGCTCGGAGCGCTTCGCCAAGTACAAACCGACTCC 1286
Qy 1355 TCGGATCGAGCAGCTGTTGGCGACGCGCGGCTCTACGCG 1396
Db 1287 TCGGATCGAGGAGGCTCGGCGCAACGCCCACTACGCCG 1328

RESULT 2
BH770842/c
LOCUS
DEFINITION
  BH770842 2491 bp DNA linear GSS 01-MAY-2002
  LJMtag581 MG1363 Random Sequence Tag Library Lactococcus lactis
  subsp. cremoris genomic, Genomic survey sequence.
ACCESSION
  BH770842
VERSION
  BH770842.1 GI:20373799
KEYWORDS
  GSS.
SOURCE
  Lactococcus lactis subsp. cremoris
  Lactococcus lactis subsp. cremoris
  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
  Lactococcus.
REFERENCE
  1 (bases 1 to 2491)
  Bolotin.A., Ehrlich.S.D. and Sorokin.A.
  Studies of genomes of dairy bacteria Lactococcus lactis
  Sci. Aliments (2002) In press
  Contact: Sorokin A
  Genetique Microbienne
  INRA
  CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
  Tel: 33 1 34 65 25 16
  Fax: 33 1 34 65 25 21
  Email: sorokine@jouy.inra.fr
  Best homologue in strain IL1403 is enoA (100%)
  Class: shotgun
  High quality sequence start: 30
  High quality sequence stop: 2463.
FEATURES
  Location/Qualifiers
    1..2491
      /organism="Lactococcus lactis subsp. cremoris"
      /mol_type="genomic DNA"
      /strain="MG1363"
      /sub_species="cremoris"
      /db_xref="taxon:1359"
      /clone_lib="MG1363 Random Sequence Tag Library"
      /note="Vector: pSGM02; Site:1; Smai; Library of
      chromosomal fragments of L.lactis strain MG1363 was
      prepared by partial AluI digestion or by sonication."

Query Match 25.2%; Score 397.6; DB 8; Length 2491;
Best Local Similarity 60.4%; Pred. No. 7.5e-36;
Matches 721; Conservative 0; Mismatches 449; Indels 24; Gaps 3;

Qy 259 GTTCCATCGCGCATCCACCGCGTCCACGAGGCTCATGAGCTGCGTGCAGTGGCGA- 317
Db 2485 GTACCTTCAGTGTCTTCTACTGTGTAACACGAGCGGTGAACCTGATGGCGACAA 2426
Qy 318 --TGTCTACTGGGCAAGGCGCTTTGAAGGCGATTGAAAACGTCAACGAAAGATCGGC 375
Db 2425 TCTCGCTACAAACGAGCTTGGTACTCAAAAAGCTTTGACAAACGTAACATCATCGCT 2366
Qy 376 GAGGAGCTCGCTGCGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAG 435
Db 2365 GAAGCTATCATCGGTTATGAATTCTGACCAACAAGCTATTGACCGTGCAATGATCGCT 2306
Qy 436 CTTGATGGCACCGCAACAAGTCCCGCTGGGTGCAACGCAATCTCTGTTGTTTCATG 495
Db 2305 CTTGACGCTACTGAAAACAAAGGTAATTTGGAGCTAACGCTATTCTTGTGTTTCTATC 2246
Qy 496 GCTGTTCAAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGA 555
```



```

Db      821  ATTGTGATCGGCATGACGGTGGCTGCTCTAGTTCTTACAAGGGCGGCAAGTACGACCTG 880
Qy      916  GAA-----GGTGGCCAGCACTCCGCGACTCCGAGCTGAGATGGCAACGTT 954
Db      881  GACTTCAAGTCTCCGAGACCGCGCGGTATCATCTCTCCGAGAGCTGGCGACCTC 940
Qy      955  TAGCTGAGTCTGTTGACGGTACCAGTACCAATCGTCTTCATGAGGAGGCCACCTGCGAGGAGAT 1014
Db      941  TACAGGGGCTTCTGTCAAAGATTACCCAGTGGTTCATCGAGGACCCCTTTGACCGAGGAC 1000
Qy      1015  GACTGGGAGGTTTACACCACTTACCGCAACCATCGGCGACAAAGTTCAGATGCTGTTGGC 1074
Db      1001  GACTGGGAGGCTGGACCAATTTTACAGCCAGCACCAAGC-----ATTAGGTGGTGGG 1054
Qy      1075  GACGACTTCTTCGTTCACCAACCTCTAGCGCTGAAGGAGGGCATCGTAAAGAGGCTGCG 1134
Db      1055  GACGACTTCAAGTCTACCGACCGGACCAAGCATCGCCAGGGTGTGGCCAGAGTCTCTGC 1114
Qy      1135  AACTCCATCTGTTTAAAGTGAACCAAGATCGGTACCTTCCATCGAGGACCTTTCGACGCTGTC 1194
Db      1115  AACTGCTGCTGCTTCAAGTCAACAGATCGGCTCGGTCCGATCCGAGTCCCTGCGAGCATGC 1174
Qy      1195  GACATGCTCACCAGGAGGATACCTCCATGATGTCGCCACCGTTCGGTGAAGACGAG 1254
Db      1175  AAGATGGCCAGAGCAACCGGTGGGGGTGATGTTGATGCTAGCCATCGCTCCGGGGAGACGGAG 1234
Qy      1255  GACACCAACCATTTGTGACCTCGCAGTTGCACTCAACTGTGTGGCCAGATCAAGACTGGTGTCT 1314
Db      1235  GACACCTTCACTGCTGACCTGGTGGTCTGTGTCACCGAGACAGATCAAGACGGGCGCT 1294
Qy      1315  CCAGCAGCTTCCGACCGTGTGCAAGATGACAAACAGCTTCTCCGATCGAGAGCTGCTT 1374
Db      1295  CCGTGGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCCGGATCGAGGAGAGCTG 1354
Qy      1375  GCGACCGCGGCTTACGAGGCTCGCAGGCGATCCAGCTTTCAAGG 1424
Db      1355  GCGACCAAGGCGCTTCCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG 1404

```

RESULT 4

```

LOCUS      CR726296              1643 bp      mRNA      linear      HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION  CR726296
VERSION    CR726296.1 GI:51224547
KEYWORDS   HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE     Tetraodon nigroviridis

```

```

ORGANISM   Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraontoidea; Tetraodontidae; Tetraodon.
            1 (bases 1 to 1643)

```

REFERENCE

```

AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
            : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            More information available at
            http://www.genoscope.cns.fr/tetraodon.

```

FEATURES

```

source
    Location/Qualifiers
        1..1643
            /organism="Tetraodon nigroviridis"
            /mol_type="mRNA"
            /db_xref="taxon:99883"
            /tissue_type="fish"

```

ORIGIN

```

Query Match      24.4%; Score 385.2; DB 3; Length 1643;
Best Local Similarity 60.8%; Pred. No. 1.5e-92;
Matches 797; Conservative 0; Mismatches 453; Indels 60; Gaps 8;

```

```

Qy      160  ATCATGCAAGTATTGCTCGCGGAAATTTCTGACTCCCGGGTAAACCCACCGTCGAGGCA 219
Db      113  ATCTCTGAAGATCCACGCTCGGAAATA TTTGACTCCCGGGCAACCCACCGTGGAGGTC 172
Qy      220  GAGGTTTTCTCGATGACGGTTCCACGCTGCGAGGTGTTCCATCCGCGGCATCCACC 279
Db      173  GACCTGTACACAAGAAAGTCTGTTTCAAGCTGG---GTCCCGAGCGCGCTCCACC 229
Qy      280  GCGCTCCACGAGGCTCATGAGCTGCGTACGCTGCGTGGCGA---TCGTTACTCTGGGCAAGGCG 336
Db      230  GGCATCTACGAGGCGCTTGGAGCTCCGCAACAACGCAAAACCCGCTTACATGCGCAAGGT 289
Qy      337  GTTTTGAAGGCAATTGAACACGTCACAGAGAAATCGGCGACAGCTCCTGCGCTTAGAG 396
Db      290  GTCTCTAAAGCTGTTGAGAATATCAATAAAATTAATTTGCACCTGSCACTGTTGGCAAGGAC 349
Qy      397  GCTGACGATCAGCGC-----TCATCGACGAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db      350  GTCAGGTTCTGGACGAGCAAAATTCGACAGCTGATGCTGGAGATGACGCGACCGAC 409
Qy      451  AACAAAGTCCCGCTGGGTGCAAAACGCAATCTTTGTTGTTTCCATGCTGTTTGAAGGCT 510
Db      410  AACAAATCTAAGTTCTGGGCGCAACGCCATCTCTGGGCGTGTCTCTGCTGTGTGAAGGCT 469
Qy      511  GCTGCTGATTCGCGAGGCTCCCATGTTCCGCTACATCGGTGGACCAACGC-----563
Db      470  GGTGACGACGAGAAAGGGCGTGGCCCTCTACCGCCACATCGCGACCTGGCGGCAACCCC 529
Qy      564  --ACAGTTCTTTCAGTTCCAAATGATGAACATCATCAACGGTGGGCTCAACGCTGACTCC 621
Db      530  CAAAGTCATCTCTCCGTTCCCGCTTTCAGGTCATCAACGGGCTCCACGCGAGCAAC 589
Qy      622  GGTGTTGAGCTTCAAGAAATTCATGATCGCTTCATCGGTGCGAGACCTTCTCTGAGGCT 681
Db      590  AAGCTGGCCATGACGAGTTTATGATCTCTGCGCGTGGAGCCAGCAGCTTCAAAGAGGCC 649
Qy      682  CTCGCAACGGCGGAGGCTTACCAAGCTGATCAAGTCCGTCATCAAGGAAAGAGGC---738
Db      650  ATGCGCATCGGCGCCGAGTCTACCAACACCTGAAGAACGTCAATCAAGAGAAAGTAGGCG 709
Qy      739  ---CTGTCACCGGACTTGGCGATGAGGCGGCTTCGCTCTCTTCCGCTCGGCTCCACCGCT 795
Db      710  AAGGATGCCACCAACGTTGGGAGACGAGGAGGCTTTGCCCCCAACATCTCTGGAGAAACAAG 769
Qy      796  GAGGCTCTTGAACCTTATCGTTGAGGCAATGAGAAAGCTTGTACCCCGAGCAAGGAC 855
Db      770  GAAGCTCTGAGCTGCTGAAGAACGCGCATCGCTTAAGGCGGCTTACAC-----CGACAAG 823
Qy      856  ATCGCTCTTCTGCTGAGCGTTGCTTCTCTGAGTTCTTCAAGGACGCGCACTTACCACTTC 915
Db      824  ATTGTGATCGGATGAGCGTGGCTGGCTCGAGTTCTAAGGGCGGCAAGTAGACCTG 883
Qy      916  GAA-----GGTGGCAGCACTCCGCAACCATCGGCGCAAGGTTTCAGATGTTGGC 954
Db      884  GACTTCAAGTCTCCGACGACCCCGCGCGCTACATTTCTCCGAGCAGCTGGCGGACCTC 943
Qy      955  TAGCTGAGCTCGTTGACGGGTACCCATCGTCTTCATCGAGGACCCACTGCGAGGAGAT 1014
Db      944  TACAGGGGCTTCGTCAAGATTACCCAGTGGTGTCCATCGAGACCCCTTTTGACCGAGAC 1003
Qy      1015  GACTGGGAGGTTTACACCAACCTTCAACGCAACCATCGGCGCAAGGTTTCAGATGTTGGC 1074
Db      1004  GACTGGGAGGCTGGACCAATTTTCAAGCAGCACCAAGC-----ATTAGGTGGTGGC 1057
Qy      1075  GACGACTTCTTCGTCACCAACCTTGAGCGCTTGAAGAGGGCATTCGTTAAGAGGCTGCC 1134
Db      1058  GACGACCTTACGCTGTCACCAACCCCAACGCAATCGCAAGGCTGTGGCCAGAGTCTCTGC 1117
Qy      1135  AACTCCATCTGTTTAAAGTGAACCAAGTACGTTACCTTACCGAGACCTTTCGACGCTGTC 1194
Db      1118  AACTGCTGCTGCTCAAGATCAACAGATCGGCTCGGTCACCGAGTCCCTTGCAGGCGATGC 1177

```

Qy	1195	GACATGGCTCACCGCGGAGGCTACACCTTCATGATGTGCCACCGTTCCGGTGAGACCGAG	1254
Db	1178	AAGATGGCCCCAGAGCAATGGCTGGGGCGTGATGTGTACGCCATCGCTCCGGGAGACGGAG	1237
Qy	1255	GACACCAACCATTTGCTGACCTCGCAGTTTGCACTCAACTGTGCGCCAGATCAAGACTGGTGCT	1314
Db	1238	GACACCTTCATCGTGTGACCTGGTGTGGTCTGCGCCAGCGACAGATCAAGACGGGCGCT	1297
Qy	1315	CCAGCAGCTTCCGACCGCTGTGCGAAAAGTACAACAGCTTCTTCGCGATCGAGCAGCTGCTT	1374
Db	1298	CCGTGCGCTCCGAGCGCTTGGCCAAAGTACAACAGCTGCTTCGCGATCGAGGAAGCTG	1357
Qy	1375	GGCGACCGCGCGTCTACGCGAGTCTGAGCGCATTTCCAGCTTTTCAGGG	1424
Db	1358	GGCGACGACGCCCGCTTCGCGCGCAAGAACTTCAGGCAACCCCATCTGAGCG	1407
RESULT 5			
BM3211196			
LOCUS	BM3211196	1505 bp	mRNA
DEFINITION	rockefeller.0.899 Mastigamoeba balamuthi lambda ZAP II Library		EST 03-JAN-2002
	Mastigamoeba balamuthi cDNA similar to enolase (EC 4.2.1.1), mRNA		
sequence.			
ACCESSION	BM3211196	1 GI:18055602	
VERSION	BM3211196		
KEYWORDS	EST.		
SOURCE	Mastigamoeba balamuthi		
ORGANISM	Mastigamoeba balamuthi		
REFERENCE	1 (bases 1 to 1505)		
AUTHORS	Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.		
TITLE	The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)		
MEDLINE	21819461		
PUBMED	11830664		
COMMENT	Contact: Muller Miklos Laboratory of Biochemical Parasitology The Rockefeller University 1230 York Avenue, New York, NY 10021, USA Email: mmuller@rockvax.rockefeller.edu Insert Length: 1505 Std Error: 0.00 POLYA=Yes.		
FEATURES			
source	1. .1505		
	/organism="Mastigamoeba balamuthi"		
	/mol_type="mRNA"		
	/strain="ATCC 30984"		
	/db_xref="taxon:108607"		
	/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"		
	/note="syn: Phreatamoeba balamuthi"		
ORIGIN			
Query Match	24.3%	Score 384.2; DB 4; Length 1505;	
Best Local Similarity	60.0%	Pred. No. 2.7e-92;	
Matches	782; Conservative	0; Mismatches 468; Indels 53; Gaps 7;	
Qy	143	AGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGCGAAATTTCTCGACTCCGCGGTA	202
Db	27	AAGCAACAATGTCGACCATCAAGAGCGTCTTCGCGCGGAGANCTTGGACTCTCGTGSCA	86
Qy	203	ACCCAAACGTCGAGCAGAGGTTTTCTTGGATGACGGTTCACCGGTGTCGAGGTGTC	262
Db	87	ACCCACCGTCGAGGTTGACCTGACCAACCGAAGAGGCGCTTTCGCGCGCGTCC	143
Qy	263	CATCCGCGGCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACGGTGGCA---TC	319
Db	144	CTCTGCTGCTCGACTGGGATCTATGAGGCGCTCGAGCTCCGTCGCGCAAGTCGC	203
Qy	320	GCTACCTGGGCAAGGGCGTTTTTGAAGCGAGTTGAAAAACGTCACCAAGAAATCGGCGACG	379

```
RESULT 6
CR679338
LOCUS          CR679338          1623 bp      mRNA          linear      HTC 19-AUG-2004
DEFINITION    Tetraodon nigroviridis full-length cDNA.
ACCESSION    CR679338
VERSION      CR679338.1 GI:51177245
KEYWORDS     HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE       Tetraodon nigroviridis
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE    1 (bases 1 to 1623)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
              : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              (E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
              More information available at
              http://www.genoscope.cns.fr/tetraodon.
FEATURES     Location/Qualifiers
              1..1623
              /organism="Tetraodon nigroviridis"
              /mol_type="mRNA"
              /db_xref="taxon:99883"
              /tissue_type="Muscle"
ORIGIN
Query Match          24.2%; Score 382; DB 3; Length 1623;
Best Local Similarity 60.7%; Pred. No. 1.1e-91;
Matches 795; Conservative 0; Mismatches 455; Indels 60; Gaps 8;
Qy 160 ATCATGACGATTTCGGTCGGGAAATTCACATCCCGCGGTAAACCAACCGTCGAGGCA 219
Db 87 ATCTGGAAGATCCACGCTCGGGAATATTCACATCCCGCGGCAACCCACCGTGGAGTC 146
Qy 220 GAGGTTTTCTCGGATGACGGTTCACCGGTGTCCGAGGTGTTCATCCCGCGCATCCACC 279
Db 147 GACTGTACACCAAGAGTCTGTTAGAGCTGG---GTCCCGAGCGGCGCTCCACC 203
Qy 280 GCGGTCCACGAGCTCATGAGTCGGTGAACCGGTGGCGA---TCGCTACCTGGGCAAGGCG 336
Db 204 GGCATCTACGAGGCCCTGGAGCTCCGCGACAAACGACAAACCCGCTACATGGGCAAGGT 263
Qy 337 GTTTTGAAGGAGTTGAAACGTCACCAAGAAATTCGGCGACGAGCTCGCTGGGCTAGAG 396
Db 264 GTCTCTAAGGCTGTTGAGAATATCAATAAATAATTGCACCTGCACTGTTGGCAAGGAC 323
Qy 397 GGTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db 324 GTACAGCTTCTGGACAGGCAAAATCGCAAGCTGATGCTGGAGATGACGGCACCGAC 383
Qy 451 AACAAATCCCGCTCGGTCGAAACCGCAATCTTGTGTTGTTCCATGCTGTTGCAAAAGGCT 510
Db 384 AACAAATCTAAGTTCCGGGCGCAACGCCATCTCTGGGCGTGTCCCTGGCTGTGTGCAAGGCT 443
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTTCGCTACATCGGTGGACCAAAACG----- 563
Db 444 GGTGACGACGAGAAGGGGTGGCCCTCTACCCGCCACATCGCGACTGGCGGCAACCCC 503
Qy 564 --ACAGCTTCTTCCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCC 621
Db 504 CAAGTCATCTCCCGCTTCCCGCTTTCAAGGTTCATCAACGGGCGCTCCACGACGCAAC 563
Qy 622 GGTGTTGAGTTGAGAAATTCATGATCGCTCCAATCGGTGCGAGACCTTCTGAGGCT 681
Db 564 AAGCTGGCCATGACGAGTTATGATCTCTGCGCGTGGGAGCCAGAGCTTCAAGAGGCC 623
Qy 682 CTCGCAACGGCGCGAGGTCTACCAACGCACTGAAAGTCCGTTCATCAAGGAAAGGCG--- 738

624 ATGCGCATCGGCGCGAAGTCTACCAACACCTGAGAGCTCATCAAGAGAGTACGCG 683
739 ---CTGTCACCGGACTTGGCGATGAGGCGGCTTGGCTCTTCCGTCGGCTCCACCGT 795
684 AAGGACGCCACCAACGTGGGAGACGAGGAGGCTTTGCCCCCAACATCTGGAGAACAG 743
796 GAGGCTCTTGACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTACCCCGCAGGAC 855
744 GAAGCTCTGGAGCTGCTGAAGAACGCCATCGGTAAAGCGCGGCTACAC-----CGAACAG 797
856 ATCGCTCTTGTCTTGACCGTGTCTTCTCTGAGTTCTTCAAGGACGCGACCTTACCACTTC 915
798 ATTGTGATCGGATGACGCTGCTGCTCGAGTCTTACAAGGCGCGCAAGTACGACCTG 857
916 GAA-----GCTGGCGAGCACTCCGCGAGCTCGAGTGAAGTGAAGAAAGCTT 954
858 GACTTCAAGTCTCCCGACGACCGCGCGCTACATCTCTCCGAGAGCTGGCGGACCTC 917
955 TAGGCTGAGCTCGTTGACCGGTACCAATCGTCTCCATCGAGGACCCACTTGCAGGAAGAT 1014
918 TACAGGGGCTTCGTCAAAAGATTACCCAGTGGTGTCTCATGAGGACCCCTTTGACGAGGAC 977
1015 GACTGGGAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAAGTTTCAGATCGTTGGC 1074
978 GACTGGGAGGCGTGGACCAATTTCAAGCCAGCACCAGC-----ATTGAGTGGTGGGC 1031
1075 GACGACTTCTTCTGTCACCAACCTGAGCGGCTGAAGAGGCGCATCGTAAAGAGGTGCC 1134
1032 GAACGCTTACGGTCAACCAACCGCATCGCAAGGGTGTGCGCCAGAAAGTCTCTGC 1091
1135 AACTCCATCTGTTAAGTGAACGAGATCGTACCTCACCGAGACCTTCGACGCTGTC 1194
1092 AACTGCTGCTGCTCAAGTCAACAGATCGGCTCGGTACCGGTACCGGTCCCTGCGAGGATGC 1151
1195 GACATGGCTCACCGCGCAGGCTACACTCCATGATGTCCACCGTTCGCGTGAAGACGAG 1254
1152 AAGTGGCCAGAGCAACCGCTGGGCGTGTGTCAGCCATCGTCCGGGAGAGACGAG 1211
1255 GACACACCATTTGTCGACCTGCGAGTTGCACTCAACTGTGCGGCGAGATCAAGTGGTGT 1314
1212 GACACCTTCACTGCTGACCTGGTGGTGGTGTGTCACCGGACAGATCAAGACGGGCGCT 1271
1315 CCAGCAGTTCCGACCGGTGCGAAAGTACAAACAGCTTCTCCGATCGAGACGCTGCTT 1374
1272 CCGTGGCTCCGAGCGCTTGGCCAGTACAAACAGCTGCTCCGATCGAGGAGAGCTG 1331
1375 GCGACCGCGCGCTTACGCGAGGTGCGAGCGCATTCACCGCTTTCAGGG 1424
1332 GCGACCGAGCGCGCTTCCCGCGCAAGAACTTCAGGACCCCATCTGAGCG 1381

RESULT 7
CR698859
LOCUS          CR698859          1613 bp      mRNA          linear      HTC 19-AUG-2004
DEFINITION    Tetraodon nigroviridis full-length cDNA.
ACCESSION    CR698859
VERSION      CR698859.1 GI:51196768
KEYWORDS     HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE       Tetraodon nigroviridis
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE    1 (bases 1 to 1613)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
              : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              (E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
              More information available at
```

[illegible]

```
Db 198 GGCACTTACGAGGCCCTGGAGCTCCGGCAACAACGACAAAACCCGCTACATGCGCAAAAGGT 257
Qy 337 GTTTTGAAGCAGTTGAAACGTTCAACGAGAAATCGGCGACAGAGCTCGCTGCGCTAGAG 396
Db 258 GTCTCTAAAGCTGTGTGAGATATCAATPAATAATATGACCTGCACCTGTGGCAAGGAC 317
Qy 397 GCTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGATGGCACC GCC 450
Db 318 GTCAACGTTCTGGACACAGGCAAAATCGAAGCTGATGTGAGATGGACGGACCGAC 377
Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATGGCTGTGTGCAAAAGGCT 510
Db 378 AACAAATCTAAGTTTCGGGGCCACGCCATCTCTGGCGTGTCTCTGGCTGTGTGCAAGGCT 437
Qy 511 GCTGCTGATTCGAGCGCTCCCACTGTTCGGCTAGATCGGTGGTGGACCAACGCC----- 563
Db 438 GGTGCAGCAGAAAGGCGTGCACCTCTACCGCCACATCCCGACCTCGGCGGCAACCCC 497
Qy 564 --ACACGTTCTCCAGTTTCCAAATGATGAACATCAATCAACGTTGGCGCTCACGCTGACTCC 621
Db 498 CAAGTCACTCTCCCGTTCCGCTTTCAAGTCAATCAACGGCGCTCCGACGAGGAAC 557
Qy 622 GGTGTTGACGTTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCT 681
Db 558 AAGCTGGCCATGCAAGGATTCATGATCCTGCGGTGGGAGCCAGCAGCTTCAAAGAGGCC 617
Qy 682 CTCCGAAACGGCGGAGGTTTACCAGCACTGAAGTCCGTCATCAAGGAAGAGGCC--- 738
Db 618 ATGCGCATCGCGCGCAAGTCTACCAAACTGAAGAACGTCATCAAGGAGGAATGACGGC 677
Qy 739 ---CTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCTCTCCGTCGGCTCCACCCTG 795
Db 678 AAGGACGCCCAACGCTGGGAGAGAGGAGGCTTTGCCCCCAACATCTCTGGAGAACAG 737
Qy 796 GAGCCTCTTGACCTTATCGTTGAGGCAATCAGAAAGCTGGCTTCAACCCAGGCAAGGAC 855
Db 738 GAAGCTCTGGAGCTGTGAAGAACGCCATCGCTAAGGCGCGCTACAC-----CGACAG 791
Qy 856 ATCGCTCTTGCTCTGGAGGTTGCTTCTCTGAGTCTTCAAGGACGCACTACCACTTC 915
Db 792 ATTGTGATCGGCATGGACGTGGCTTGCCTCTGAGTCTTCAAGGGCGGCAAGTACGACCTG 851
Qy 916 GAA-----GGTGGCCAGACACTCCGACGCTGAGATGGCAAAAGCTT 954
Db 852 GACTTCAAGTCTCCGAGACCCCGCGCTACATCTCTCCGAGAGCTGGCCGACCTC 911
Qy 955 TACGCTGAGCTGTGTGAGGTACCCCAATCGTCTCCATCGAGGACCCACTGCAAGAAAT 1014
Db 912 TACAGGGGCTTCGTCAAAGATTACCCAGTGGTGTCCATCGAGGACCCCTTTGACGAGGAC 971
Qy 1015 GACTGGGAGGTTACACCACTTACCGCAACCATCGGCGACAGAGGTTCAAGTCTGTTGGC 1074
Db 972 GACTGGGAGGCGTGAGCAATTTTACAGCCAGCACCCAGC-----ATTGAGGTGGTGGGC 1025
Qy 1075 GACGACTTCTTCGTCAACCAA-CCCTGAGCGCCTGAAGAGGGGCTCGCTAAGAAAGGCTGC 1133
Db 1026 GACGACCTCAGGTTCAACCAACCCCAACGATCGCCAAAGGTTGCGCCAGAGTCTCTG 1085
Qy 1134 CAATCCATCTCGTTAAGGTGAACCAAGATCGGTACCCCTCAACGAGACCTTCGACGCTGT 1193
Db 1086 CAATGCTGCTGTGCTCAAAGTCAACAGATCGGTCTCGGTCAACGAGTCCCTTCGAGGCAATG 1145
Qy 1194 CGCATGGCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTTCCGGTGAGACCGA 1253
Db 1146 CAAGATGGCCAGAGCAACCGGCTGGGGCGTGATGTGAGCCATCGCTCCGGGAGACCGA 1205
Qy 1254 GGACACCAACCTTGTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGC 1313
Db 1206 GGACACCTTTCATCGCTGACCTGGTGTGCTGTGCAACCGGACAGATCAAGACGCGGCGC 1265
Qy 1314 TCCAGCAGTTCCGACCGGTGTGCAAGTACAACGAGCTTCTCCGCATCGAGCAGCTGCT 1373
Db 1266 TCCGTGGCGCTCCGAGCGCTTTGGCCAAAGTACAACGAGCTGTCTCCGATCGAGGAAGACT 1325
```

```
Qy 1374 TGSCGAGCGCGCGTCTACGAGTGCAGGCGCATTTCCACGCTTTTCAGGG 1424
Db 1326 GGGCGACCAAGGCCGTTCCCGCAAGAACTTCAGGACCCCATCTGAGGG 1376

RESULT 9
CR728153
LOCUS CR728153 1637 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR728153
VERSION GI:51226414
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1637)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(BE-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1637
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

ORIGIN
Query Match 23.9%; Score 377; DB 3; Length 1637;
Best Local Similarity 60.4%; Pred. No. 2.5e-90;
Matches 791; Conservative 0; Mismatches 460; Indels 58; Gaps 8;

Qy 160 ATCATGACGATATTCGTCGCGAAATTCGACTCCCGGGTAACCCACGTCGAGGCA 219
Db 108 ATCTGAAGATCCACGCTCGGAAATATTCGACTCCCGCGCAACCCACGTCGAGGTC 167
Qy 220 GAGGTTTCTCGATGACGGTCCACGGTGTCCAGGTGTTCCATCCGCGCATCCACC 279
Db 168 GACTGTACACCAAGAAAGGTCTGTACAGAGTGTGGTCC--CCAGCGCGGCTCCACC 225
Qy 280 GGGCTCCACGAGGCTCATGAGCT--GCGTGACGGTGGCGATCGCTACTGGGCAAGGGCG 337
Db 226 GGCATCTACGAGGCCCTGGAGCTCCGCAACGACAAACCCGCTACATGGGCAAGGTG 285
Qy 338 TTTTGAAGGAGTTGAAACGTCGCAAGAAATCGCGAGAGCTCGTGGCTAGAGG 397
Db 286 TCTCTAAGCTGTGTGAATATCAATAAAATAATTGCACCTGCATCGTGTGGCAAGGACG 345
Qy 398 CTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTTGATGGCAACGCCA 451
Db 346 TCAACGTTCTGGACCAAGGCAAAATCGAAGCTGATGCTGGAGATGGACGCGACCGACA 405
Qy 452 ACAAGTCCCGCTGGGTGCAAAAGCAATCTTGGTGTTCATGGCTGTGTGCAAGGCTG 511
Db 406 ACAATCTAAGTACGGGGCCAAAGCCATCTCTGGCGTGTCTCTGGCTGTGTGCAAGGCTG 465
Qy 512 CTGCTGATTCCGAGGCTCCCACTGTTCGCTACATCGGTGGACCAACGCC----- 563
Db 466 GTGACGAGAGAGGGGCTGCCCTCTTACCGCCACATCGCCGACCTGGCGGGAACCCCC 525
Qy 564 -ACACGTTCTTCCAGTTTCCAAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCG 622
Db 526 AAGTCATCTCTCCCGTTCCCGCTTTCAACGCTCATCAACGGCGGCTCCACGACGGAACA 585
Qy 623 GTGTTGAGGTTTCAGGAATTCATGATCGCTCAATCGGTGCAGAGACCTTCTCTGAGGCTC 682
```

```
Db 586 AGCTGGCCATCAGGAGTTCATGATCTCTGCCGCTGGGAGCCAGCAGCTTCAAGAGGCCA 645
Qy 683 TCCGCAACCGCGCGAGGTTCTACCAACGACCTGAAGTCCGTCTCATCAAGGAAAGGCG--- 738
Db 646 TGCAGCATCGGCGCGAAGTCTACCAACCTGAAGAGCTCATCAAGGAGGATACGGCA 705
Qy 739 --CTGTCCACCGGACTTGGCGATGAGGCGCTTTCGTCTCTTCCGTCCGCTCCACCCGTG 796
Db 706 AGGACGCCACCAACGCTGGGAGACGAGGAGGCTTTGCCCCCAACATCTCGGAGAACAGG 765
Qy 797 AGGCTCTTGACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTCAACCCAGGCAAGGACA 856
Db 766 AAGCTCTGGAGCTCTGAAGAACGCCATCGCTAAGGCGCGGTACAC-----CGACAAGA 819
Qy 857 TCGCTCTTGCTCTCGACCTTCTCTGAGTCTCTCAAGGAGCGGACCTTACACATTCG 916
Db 820 TTGTGATCGGATGAGCTGTCTCGCTCCGAGTCTTCAAGAGGCGGCAAGTACGACCTGG 879
Qy 917 AA-----GGTGGCCAGCACTCCGCACTCGGAGCTGAGATGCAACAGTTT 955
Db 880 ACTTCAAGTCTCCGAGACGCCGCGCTACATTTCTCCGAGCAGCTGSCCGACCTCT 939
Qy 956 ACGCTGAGCTGTTGACGCGTACCAATCGTCTCTCAATCGAGGACCCACTCGAGGAAGTG 1015
Db 940 ACAGGGGCTTCTCAAGATTAACCCAGTGTGTCTCATCGAGGAGCCCTTTGACAGGAGC 999
Qy 1016 ACTGGGAGGTTACACCAACTCTACCGCAACCATCGGCGACAGGTTTCAGATCTTGGCG 1075
Db 1000 ACTGGGAGGCTGAGCAATTTTACGCCAGCACGAGC-----ATTGAGTGTGGCG 1053
Qy 1076 ACAGCTTCTTCTTCCACCAACCTGAGCGCTGGAAGGAGGCGATCGCTAAGAGGCTGCA 1135
Db 1054 ACGACTTACGCTTACCAACCCCAACGACGCGCAAGGTTGTGCCAGAGTCTCTGCA 1113
Qy 1136 ACTCATCTCTGTTAAGGTGAACAGATCGGTAACCTCACGAGACCTTTCGACGCTGTG 1195
Db 1114 ACTGCTCTGCTCTCAAGTCAACAGATCGGCTCGTCAACGAGTCTCCCTGCGAGCATGCA 1173
Qy 1196 ACATGCTCACCGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGAGACCGAGG 1255
Db 1174 AGATGGGCCAGAGCAACGCGTGGGCGGTGATGTCAGCATCTCGTCCGGGAGAGCGAGG 1233
Qy 1256 ACACACCATGCTGACCTCGCAGTTGACACTCAACTGTGGCCAGATCAAGACTGTGCTC 1315
Db 1234 ACACCTTATCGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1293
Qy 1316 CAGCACCTTCGACCGCTGTGCAAGATCAACACGCTTCTCCGATCGAGCAGCTGCTTG 1375
Db 1294 CGTCCGCTCGAGCGCTTGGCCAAAGTACAAACGAGTCTCCGATCGAGGAGAGCTGG 1353
Qy 1376 GCGAGCGGGGCTTACGAGGTGCGAGGATTCACCGATTCACCGCTTTCAGGG 1424
Db 1354 GCGACAGGCGCGTTCGCGGCAAGAACTTCAGGACCCCATCTGAGCG 1402
```

RESULT 10

```
CR685932 LOCUS 1551 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR685932
VERSION CR685932.1 GI:51183839
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage
```

```
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source 1.1551
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Eye"
```

ORIGIN

```
Query Match 23.4%; Score 369.6; DB 3; Length 1551;
Best Local Similarity 60.3%; Pred. No. 2.4e-88;
Matches 782; Conservative 0; Mismatches 454; Indels 60; Gaps 8;

Qy 174 CGCTCGGAAATTCGACTCCCGCGTAAACCCAGCTCGAGGAGAGGTTTCTCTGGA 233
Db 27 CGCTCGGAAATATTCGACTCCCGCGCAACCCACCGTGGAGGTGCACTGTACACCA 86
Qy 234 TGACGGTTCCACGCTGTCGAGGTGTTCCATCCGCGCATCCACCGCGTCCAGAGGC 293
Db 87 GAAAGGTCTGTTAGAGCTGCG---GTCCCGAGCGCGCTCCACCGGCATCTACGAGC 143
Qy 294 TCATGAGCTGCGTG---ACGGTGGCGATCGCTACCTGGCAAGGGCGTTTGAAGGCAGT 350
Db 144 CTTGAGCTCGCGACAAAGATTAAACCCGCTACATGGCAAGGGGTCAAAAGGCGCT 203
Qy 351 TGAAGCGTCAAGAGAAATCGCGGACGAGCTCGCTGGGCTAGAGGCTGACGATCAGCG 410
Db 204 TAAATATATCAATGAGTTTTGGCCCCCGCTTGTGTAACCGAGGACGTCAACGTTCTGGA 263
Qy 411 CC-----TCATCGACGAAGCATGATCAAGCTTGATGGGACCGCCCAACAGTCCCGCT 464
Db 264 CCAGGCAAAATCGACAAGCTGATGCTGGAGATGACGCGCACCAACAACTTAAGTT 323
Qy 465 GGGTGAAGAAAGCAATCTGCTGTTTCCATGGCTGTTGCAAGGCTGCTGCTGATTCCGC 524
Db 324 CGGGGCCAAGCCCATCTGGCGGTGTCCTGGCTGTGTGCAAGGCTGGTGAGAGAGAA 383
Qy 525 AGGCTTCCCACTGTTCCGCTACATCGGTGACCAACAG-----CACACGTTCTTCC 575
Db 384 GGGCGTCCCTCTTACCGCCACATTTGCGACCTTGGCGGCAACCCCAAGTCATCTCCC 443
Qy 576 AGTTCAATGATGAACATCATCAAGGTGGGCTCAGCTGACTCCGGTGTGAGTTCA 635
Db 444 CGTTCCGCTTTCAACGCTGATCAACGCGGCTCCACGCGAGGCAACAGCTGGCATGCA 503
Qy 636 GGAATTCATGATCGCTCCAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGC 695
Db 504 GAGTTTATGATCTGCGGTGGAGCAGCAGCTTCAAGAGGCGCATTCGCATCGGGCGC 563
Qy 696 GAGGCTTACCAACCACTGAAAGTCCGCTCATCAAGAGAAAGGGC-----CTGTCCACCGG 749
Db 564 CGAGGTCTACCAACCACTGAAGACGTCATCAAGAGAGATGACGCAAGGACGCCACCA 623
Qy 750 ACTTGGGATGAGGCGGCTTCCGCTTCCGTCGGCTCCACCGCTGAGGCTCTTGACCT 809
Db 624 CGTGGAGAGAGAGGAGGCTTTGCCCCCAACATCTCTGGAGAAACAGGAAGCTCTGGAGCT 683
Qy 810 TATCGTTGAGCAATCGAGAGGCTGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCT 869
Db 684 GCTGAAGAAAGCCATCTGCTAAGGCGGCTACAC-----CGACAAGATTGTGATCGGAT 737
Qy 870 GAGCGTTGCTTCTCTGAGTCTTCAAGAGCGGCACTACACATTCGAA----- 918
Db 738 GAGCGTGGCTGCTTCCGAGTTCTCAAGGGGCGCAAGTACGACTGACTTCAAGTCCCC 797
Qy 919 -----GGTGGCCAGCATCTCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTCGT 968
Db 798 CGACGAGCCCGCGCTTACATCTCTCCGAGCAGCTGCGCGACCTCTACAGGGGCTTCGT 857
```



```
Qy 1313 CTCGACGACGTTCCGACCGTGTCCGAAAGTACACAGCTTCTCCGATCGAGCAGTGC 1372
|||||
Db 1293 CTCGTCGCGCTCCGAGCGTGTGGCCAAAGTACACAGCTGCTCCGATCGAGGAGC 1352
|||||
Qy 1373 TTGGCGACCGCGCGTCTACCGAGTGCAGCGCATTCCTCCAGCTTTTCAGGG 1424
|||||
Db 1353 TGGCGACACAGCGCCCTTTCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG 1404

RESULT 12
CR656482 1582 bp mRNA linear HTC 18-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR656482
ACCESSION CR656482
VERSION CR656482.1 GI:51152927
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1582)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source
1. 1582
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Query Match 23.1%; Score 364.6; DB 3; Length 1582;
Best Local Similarity 60.3%; Pred. No. 5.5e-87;
Matches 791; Conservative 0; Mismatches 459; Indels 61; Gaps 9;

Qy 160 ATCATGACGATTTTCGGTCCGGAATCTCCAGCTCCGCGGTAAACCAACCGTCGAGCA 219
|||||
Db 67 ATCTGGAATCCAGCTCGGAAATATTGACATCCCGCGCAACCCACCGTCGAGGTC 126
|||||
Qy 220 GAGTTTTCTGGATGACGGTTCACCGGTGTCGAGGTGTTCATCCGCGGCATCCACC 279
|||||
Db 127 GACCTGTACACCAAGAAAGTCTGTTACAGAGCTGCG--GTCCCGACGCGCGCTCCACC 183
|||||
Qy 280 GCGTTCACGAGGCTCATGAGCTCGGTGACGGTGGCGA---TCGCTACCTGGGCAAGGC 336
|||||
Db 184 GGCATCTACGAGGCGCTGGAGCTCCGCGACAAACCAACCAACCCGCTACATGGGCAAGGT 243
|||||
Qy 337 GTTTTGAAGCGAGTTGAAAGCTCAACGAAGAAATCCGCGACGAGCTCGTGGCCTAGAG 396
|||||
Db 244 GTCTTAAGCTGTGTGAATATCAATAAATAATATGACCTGCACCTGGTTGGCAAGGAC 303
|||||
Qy 397 GCTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCAGCGCC 450
|||||
Db 304 GTCAACGTTCTGGACAGGCAAAATTCGCAAGCTGATGCTGGAGATGGACGGCACCGAC 363
|||||
Qy 451 AACAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCGCTGTGTCGAAAGGCT 510
|||||
Db 364 AACAAATCTAAACACGGGCGCAACGCCATCTCGGCGGTGTCCCTGGCTGTGTGCAAGGCT 423
|||||
Qy 511 GCTGCTATTCCGCGAGCGCTCCCACTGTTCGCTTACATCGGTGACCAAGC----- 563
|||||
Db 424 GGTGACAGAGAGGGGTGCGCACTTACCGCCACATCGCCGACCTGGCGCGCAACCC 483
|||||
Qy 564 --ACACGTTTTCAGGTTTCCAATGATGAACATCATCAACGGTGGCGGTCAAGCTGACTCC 621
|||||
```

```
Db 484 CAAGTCATCTCCCGCTTCCCGCTTTCAACGTCATCAACGCGCGCTCCACGCAAGAAAC 543
|||||
Qy 622 GGTGTTACGCTTCAGGAATTCATGATCGCTCAATCGGTGCGAGAGACTTCTCTGAGGCT 681
|||||
Db 544 AAGCTGCGCATGCGAGGAGTTCATGATCTCTCGGTGGGAGCCAGCAGCTTTCAAAGAGGCC 603
|||||
Qy 682 CTCCGCAACGCGCGGAGGCTTACACGACACTGAAGTCCGTCATCAAGAGAAAGGCG--- 738
|||||
Db 604 ATGCGCATCGCGCGGAGTCTACCAACACTGAAGAACGTCATCAAGGAGAGATACGGC 663
|||||
Qy 739 ---CTGTCTCACCGGACTTTGGCGATGAGGGCGCTTCGCTCTCTCGTGGGTCCACCGCT 795
|||||
Db 664 AAGGACGCCCAACGCTGGGAGACGAGGAGGCTTTGCCCCCAACATCTCTGGAGAACAG 723
|||||
Qy 796 GAGGCTCTTGACCTTATCGTTGAGGCAATCGAAGAGCTGGCTTACCCCCAGGCAAGGAC 855
|||||
Db 724 GAAGCCCTGGAGCTGCTGAAGAACGCTCGTGAAGCGCGGTACAC-----CGACAAG 777
|||||
Qy 856 ATCGCTCTGCTCGGACGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTACCACCTTC 915
|||||
Db 778 ATTGTGATCGCATGGAGCTGGCTGCTCTGAGTTCTACAAGGGCGGCAAGTACGACCTG 837
|||||
Qy 916 GAA-----GGTGGCGAGCACTCCGCGAGCTCCGAGCTGAGATGGCAAAAGCTT 954
|||||
Db 838 GACTTCAAGTCTCCCGACGACCCCGCGCTTACATCTCTCCGAGCAGCTGGCGACCTC 897
|||||
Qy 955 TAGCTGAGCTCGTTGAGCGGTACCCCAATCGTCTCAATCGAGGACCCACTTCGAGGAGAT 1014
|||||
Db 898 TACAGGGGCTTCGTCGCAAAAGATTACCCAGTGGTGTTCATCGAGGACCCCTTTTGAC 957
|||||
Qy 1015 GACTGGAGGAGGTTACACCAACCTCACGCAACCAATCCGCGACAAAGGTTTCAGATCGTTGGC 1074
|||||
Db 958 GACTGGAGGAGGTTGACCAATTTTCAAGCAGCAGCAGC-----ATTGAGTGGTGGC 1011
|||||
Qy 1075 GAGGACTTCTTCGTCACCA--CCGTGAGCGCTGAAGAGGAGGATCGCTAAGAGGCTGC 1133
|||||
Db 1012 GAACGCTCTACGGTCAACCAACCCCAACGATCGCCAAAGGCTGTGGCCCAAGAGTCTCTG 1071
|||||
Qy 1134 CAATCCATCTGTTAAGGTGACGAGTCCGTACCTCCTCACGAGACCTTCGAGCGCTGT 1193
|||||
Db 1072 CAATGCTCTGCTGCTCAAGATCAACAGATCGGCTCGGTCAACCGAGTCTCTGCGAGGATG 1131
|||||
Qy 1194 CGACATGGCTCACCGCGAGCTACACCTCATGATGTCCACCGTTCGCGTGAGACCGA 1253
|||||
Db 1132 CAAGATGGCCGAGAGCAACGCTGGGCGGTGATGGTCAGCCATCGCTCCGGGAGAGCGGA 1191
|||||
Qy 1254 GGACACCACTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGCTGTC 1313
|||||
Db 1192 GGACACCTTCATCGTTCTCTCTGTTGGTGGTCTGTGCAACCGACAGATCAAGACGGGCGC 1251
|||||
Qy 1314 TCCAGCAGCTTCGACCGTGTGCAAGATCAACAGCTTCTCCGATCGAGCAGCTGCT 1373
|||||
Db 1252 TCCGTGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCCGGATCGAGGAGAGCT 1311
|||||
Qy 1374 TGGCGACGCGGGCTCTACGAGCTCGCAGCGATTCGCCAGCTTCCACGCTTTCAGGG 1424
|||||
Db 1312 GGGCGACGAGCGCGTTCGCGGCAAGAACTTCAGGCACCCCATCTGAGCG 1362
|||||

RESULT 13
BH770828 2021 bp DNA linear GSS 01-MAY-2002
LOCUS LMGag569 MG1363 Random Sequence Tag Library Lactococcus lactis
DEFINITION subsp. cremoris genomic, genomic survey sequence.
ACCESSION BH770828
VERSION BH770828.1 GI:20373785
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 2021)
```

AUTHORS Bolotin, A., Ehrlich, S. D. and Sorokin, A.
 TITLE Studies of genomes of dairy bacteria *Lactococcus lactis*
 JOURNAL Sci. Alimentis (2002) in press
 COMMENT Contact: Sorokin A
 Genetique Microbienne
 INRA
 CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
 Tel: 33 1 34 65 25 16
 Fax: 33 1 34 65 25 21
 Email: sorokine@jouy.inra.fr
 best homologue in strain IL1403 is enoB (95%)
 Class: shotgun
 High quality sequence start: 30
 High quality sequence stop: 1993.
 Location/Qualifiers
 1. .2021
 /organism="Lactococcus lactis subsp. cremoris"
 /mol_type="genomic DNA"
 /strain="MG1363"
 /sub_species="cremoris"
 /db_xref="taxon:1359"
 /clone_lib="MG1363 Random Sequence Tag Library"
 /note="Vector: pSGM2; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MG1363 was prepared by partial AluI digestion or by sonication."
 ORIGIN
 Query Match 21.7%; Score 341.8; DB 8; Length 2021;
 Best Local Similarity 56.9%; Pred. No. 8.4e-81;
 Matches 691; Conservative 0; Mismatches 512; Indels 12; Gaps 3;
 QY 159 AATCATGACGTATTCGTCGGAATTCCTCGACTCCGCGGTAAACCAACCGTCGAGGC 218
 DB 709 AATCGAARACATTCACGCAAGAGAAATTTTGATTCAGAGGAAATCCAAACCGTTGAGGT 768
 QY 219 AGAGGTTTCTCGATGACGGTTCCACGGTGTCCAGGTGTTCCATCCGCGCATCCAC 278
 DB 769 AGATGTTAGATTGACAGATGAATTTGGGACGGCGGCGAGTTCCTCAGAGCATCAAC 828
 QY 279 CGGCGTCCACGAGCTCATGAGTCGCTGCGGTGCGGTGCGATCGTACCTGGGCAAGGCGGT 338
 DB 829 AGGAGATCGTGAAGCGGTGGAATTAAGAGATGGCGGAGCGGAGTTCGAAGGAAAGCGGT 888
 QY 339 TTTGAGGACGTTGAACGTTCAACGAGAAATCGGCGACGAGCTCGCTGCGCTAGAGGC 398
 DB 889 TTTCAAGCTGTTGCAATGTTAATGTTGAAATTTATGAAGCCTTTAAAGGCCAATCACC 948
 QY 399 TGAGCATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGCGACCGCCAAACAAGTC 458
 DB 949 ATTCAATCAAGCAAAATTAGACCATCTGATGATTTGAATTTGATGGACAAAATAAGTC 1008
 QY 459 CCGCTGGGTGCAACGCAATCTTGGTGTTCATGCGCTGTTGCAAAAGGTGCTGTGA 518
 DB 1009 TCGCTTAGTGCAATGCAATATTGGGTGTTTCCATGCTATTGCTGCGCGCAGCTAA 1068
 QY 519 TTCGCGAGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACACGCTTCTCCAGT 578
 DB 1069 CAGTGAAAAATCCCACTTTATCGTTATCTTGGTGGAGTAGATTGGAACTTCCTCAA -- 1126
 QY 579 TCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGCTGTTGACGTTTCAGGA 638
 DB 1127 -CCATTTTCAATGTCATCAATGTTGGATGTCACGCCGATTCAGAAATTGAGTTCAAGA 1185
 QY 639 ATTCAATGATCGTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGA 698
 DB 1186 ATTCTTGATTACCCCTGTCAACACGTGCTAGTTTTCGCGATGGTTTGAAGAAAATCGCAA 1245
 QY 699 GGTCTACCAACGACTGAAGTCGTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGA 758
 DB 1246 TATCTATCATCTTTGAAGAAAATTTTATGCTGATTAAGAGCTAGAAACGGCGGTAGGGA 1305
 QY 759 TGAGGCGGCTTCGCTTCCTTCGTCGCTCCACCGCTGAGGCTCTTGACCTTATCGTTGA 818

DB 1306 CGAGGGTGCTTTTGCTCCAAACTTGTTCCAGCGAAAAATGCCATTGCTACACTTTATCA 1365
 QY 819 GGCATTCGAGAAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTGAGACGTTGC 878
 DB 1366 AGCAATTGAAAGAGCTGGTTATGTGCGCAGCGGAAAGAAATTTGGGATTGCTATTGACCCGGC 1425
 QY 879 TTCTCTGAGTTTCTT-----CAAGGACGGCACCTACCACTTCGAAGGTGGCCGAGCACTC 932
 DB 1426 TTCTAGCGAATTTTATGATGACAAAGAAAAGTCTATCATTTTGAAGGTCAAAAGCTCAC 1485
 QY 933 CGCAGCTGAGATGGCAAAAGCTTTACGCTGAGCTGTTGACGGGTACCCAAATCG---TCTC 989
 DB 1486 TTCAGCAGAAATTTGTCACCTAAATATGAGGGTTAGTTGAAAAATATCTGCCCTTAATCTC 1545
 QY 990 CATCGAGGACCCACTGCAGGAGATCACTGGGAGGTTTACACCAACTCACCGCAACCAT 1049
 DB 1546 TATTGAAGACGGATTTTCGGAACATGATTGGGCTGGTTTTCGGGCTCAACAGAAAGTTCA 1605
 QY 1050 CGCGCAAGAGTTTCAGATCGTTGGCGACGACTTCTTCGTCCACCAACCTTGAGCGCTGAA 1109
 DB 1606 AGTCAAAAATTTCAATTTGGTGTGATGATATTTTGTACAAATCCCGAAATCTTCAA 1665
 QY 1110 GGAGGCGATCGCTAAGAGGCTGCCAACTCCATCTCTGGTTAAGGTGAACAGATCGGTAC 1169
 DB 1666 AGAAGGAATTTCAAAAAAGTGTAGCAAAATGCTATTTTGAATTAATACTAAATCAATCGGAC 1725
 QY 1170 CTTACCGGAGACCTTCGACGCTGTCACATGGCTCACCGCGGAGGTACACCTCCCATGAT 1229
 DB 1726 AGTGACCGAAGCAATTTGAAGCGATTAGTTTGGCCAGAAAAGAGGTTTATGAACCATGAT 1785
 QY 1230 GTCCCAACCGTTCCGCTGGAGACCGGACACCACTTGTCTGACCTCGCAGTTGCACTCAA 1289
 DB 1786 TTCTCATCTGTTAGGTGAAACGGTTGATAGTTTATTTGCTGACTTTGCTATAGCCATGCA 1845
 QY 1290 CTGTGCGCAGATCAAGACTGGTCTCCAGCACTCCGACCGTGTGCGCAAAAGTACAACCA 1349
 DB 1846 CGCTGCAAAATAAAACAAGTTCAATGGCGAAGTGAGCGGTTGAAAAATATATATCA 1905
 QY 1350 GCTTCTCCGATCGA 1364
 DB 1906 ATTTTGGCGATTGA 1920
 RESULT 14
 CR658829
 LOCUS Tetraodon nigroviridis full-length cdna. linear HTC 18-AUG-2004
 DEFINITION
 ACCESSION CR658829
 VERSION CR658829.1 GI:51155274
 KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 1 (bases 1 to 1551)
 REFERENCE
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
 : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 More information available at
 http://www.genoscope.cns.fr/tetraodon.
 FEATURES
 Location/Qualifiers
 1. .1551
 /organism="Tetraodon nigroviridis"
 /mol_type="mRNA"
 /db_xref="taxon:99883"
 /tissue_type="Muscle"
 ORIGIN

Query Match	20.8%; Score 328; DB 3; Length 1551;	
Best Local Similarity	57.9%; Pred. No. 4.2e-77;	
Matches	740; Conservative 0; Mismatches 490; Indels 48; Gaps 7;	
Qy	160 ATCATGCACGATTTTCGCTCGCGAAATTTCTCGACTCCCGCGTAAACCAACCGTCGAGGCA	219
Db	1098 CTGCTGCTCAAGGTCAACAGATCGGCTCAGTGACCGAGTCCATCAAGACGCTGAAGCTG	1155
Qy	87 ATCACTAAGATCCAGGCTCGGAGATCTTGACTTCAGAGAAACCCCAACGCTGGAGTG	146
Db	1201 GCTCACGCGCAGGCTACACCTCATGATGTGCCACCGTTCCGGTTCAGAGCCGAGGACACC	1360
Qy	220 GAGGTTTTCTTGGATGACGGTTCCCAAGGTGTCCAGAGGTGTTCCATCCGCGCATCCACC	279
Db	1158 GCCACAGACAGTGGTTGGGTGTGATGGTCAGCATCGCTCTCGAGAAACTGAGGACACC	1217
Qy	147 GACCTGTGGACGGCAAAGG--CCTGTTACAGACAGCGCTCCCCACGCGGGCTCTACT	203
Db	1261 ACCATTGCTGACTCCGACTCCGAGTTGCATCAACTGTGGCCAGATCAAGACTGGTCTCCAGCA	1320
Qy	280 GGCCTCCACAGGCTCATGAGCTCGGTGCGGTGGCA--TTCCTACCTGGGCAAGGC	336
Db	1218 TTCTATCCGACCTGGTGTGGACTGTGCACCGACAGATCAAGACGGGCGCTCCCTGC	1277
Qy	204 GGGTTCATGAGCGCTGGAGCTTCGTGATGGAGACAAGATGCTACTTGGCAAGGA	263
Db	1321 CGTTCGACGGTCTCCAAAGTACAACAGGCTTCTCGGATCGAGAGCTGCTTGGCGAC	1380
Qy	337 GTTTTGAAGGCAGTTGAAGACGTCAAAGAAATCGGACGAGCTGCGTG-----GC	390
Db	1278 AGGTCGGAGCGTTTGGCCAAAGTACAACAGGCTGATGAGAAATCGAGGAGCTTGGGAC	1337
Qy	264 ACCCTGAAGCCGTGGACCACTCAACAGGACATCGCCCAAGCTGATGAGAGAAA	323
Db	1381 GCCGGGCTCTACGACAGT 1398	
Qy	391 CTAGAGGCTGACGATCAGCGCCTCATCGAGCAAGCAATGATCAAGCTTGATGGCAGCC	450
Db	1338 AAGCCCAAGTTCGCCGCT 1355	
Qy	324 TTCAGCGTTGGAGCAGGAGAAATTCACAGTTTCATGCTGGAGCTGGATGGAACCGAG	383
Db	CR733357	1407 bp mRNA linear HTC 19-AUG-2004
Qy	451 AACAACTCCCGCTCGGTGCAACGAAATCTTGGTGTTCATGCTGCTGTTGCAAGGCT	510
Db	CR733357.1 GI:51231674	
Qy	384 AACAAATCTCAGTTTGGCGCAACGCCATCTTGGCGGTGTCCTTGGCGCTCTGCAAGCC	443
Db	HTC; cDNA; full-length; Tetraodon nigroviridis.	
Qy	511 GCTGCTGATTCCGACGCGCTCCCACTCTCCGCTACATCGGTGACCAACGACACG--	568
Db	Tetraodon nigroviridis	
Qy	444 GGGCGGGGAGAAAGGCTTCTCTCTACCGCACATCGCCGATCTCGCGGACACAAA	503
Db	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	569 -----TCTTCCAGTTTCCAATGATGAACATCATCAACGCGTGGCGCTCACGCTGACTCC	621
Db	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
Qy	504 GACGTGATCTCCCGTCTCTGCTTTCAACGTCAATCAATGGCGGACCAAGCTGGGAAC	563
Db	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;	
Qy	622 GGTGTTGACGTTACAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTCTGAGGCT	681
Db	Tetraodontidae; Tetraodontidae; Tetraodon.	
Qy	564 AAGCTGGCCATGACAGGATTCATGATTTCCCGCTTGGAGCGGCAACTTCCACAGGCC	623
Db	1 (bases 1 to 1407)	
Qy	682 CTCGCAACGCGCGGAGGCTACACGCACTGAAGTCCTGATCAAGGAAAGGGC---	738
Db	Genoscope.	
Qy	624 ATGAGGATCGCGCTGAGGTTTACCAACCTGAAGAGCGTGATCAAGGCAAAAGTACGGA	683
Db	Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage	
Qy	739 ---CTGTCCACCGACTTGGCGATGAGGGGGCTTCGCTCTCTCCGTCGCTCCACCCGT	795
Db	: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE	
Qy	684 AAGAGCCACCAACGTTGGAGAGAGAGGCTTCGCCCAACATCTCTGGAGAACAC	743
Db	(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
Qy	796 GAGCTCTTACCTTATCGTTGAGGCAATCGAGAGGCTGGCTTCACCCAGGCAAGGAC	855
Db	The sequences are based on single pass reads.	
Qy	744 GAAAGCTGAGGCTGTGAAGTCAGCCATCGAGAGCGGCTACCCCGCAAGATCATC	803
Db	More information available at	
Qy	856 ATGCTCTTCTCTGAGGCTTCTCTCTGAGTTCTTCAAGAGCGCACTTACCACTTC	915
Db	http://www.genoscope.cns.fr/tetraodon.	
Qy	804 ATCGGATGAGCGTGGCGCTCGAGTTCTATCGCAGCGGGAAGTACGACCTTGA	863
Db	Location/Qualifiers	
Qy	916 GA-----AGGTGGCCAGCACTCCGACGTGAGATGGCAACGTTTACGCT	960
Db	1. .1407	
Qy	864 AAGTCCCCGATGACCCATCCAGGCATACAGCGGGGAGAGCTGGAGACCTGTACCGC	923
Db	/organism="Tetraodon nigroviridis"	
Qy	961 GAGCTCGTTGACGCTACCCAACTCGTCTCCATCGAGGACCCACTGCAAGGAGATGACTGG	1020
Db	/mol_type="mRNA"	
Qy	924 AGCTTATCCAGAACTACCTGTTTCAGTTCATCGAAGACCCCTTTGACCAAGGATGACTGG	983
Db	/db_xref="taxon:99883"	
Qy	1021 GAGGTTTACACCACTCACCGCAACCATCGGCGACAAGTTTCAGTTCGTTGGGAGCAC	1080
Db	/tissue_type="fish"	
Qy	984 GAGAACTGGGCCAAGTTTACCGCTCCACAGACATC-----CAGATCGTTGGAGATGAC	1037
Db	Query Match	20.7%; Score 327; DB 3; Length 1407;
Qy	1081 TTCTTCGTCAACCACTGAGCGCTCAAGAGGGCATCGTAAGAGGGCTGCCAACTCC	1140
Db	Best Local Similarity	59.7%; Pred. No. 7.6e-77;
Qy	1038 CTAGCGGTGACCAACCCCAAGAGATCCAGAGCGGTTCGACAGAGGCTGTAACTGT	1097
Db	Matches	686; Conservative 0; Mismatches 410; Indels 53; Gaps 6;
Qy	1141 ATCCTGGTTAAGGTGAACAGATCGGTATCCCTCACCGAGACCTTTCGACGCTGCGACATG	1200
Db	317 ATGCTACCTCGGCAAGGCGCTTTTGAAGCGAGTTGAAACCGTCAACGA-----AGAAAT	371
Qy	372 CGGCGAGAGCTCGCTGGCTGAGAGCTGACGATCAGCGCTCATCGACGAAATGAT	431
Db	13 ATGGCTACTGGGCAAGGGGTCARAAAGGCGCTTAATATGTCAATGAGTTTTTGCCCC	72
Qy	73 CGCCTTGTGTAACAGGAGCTCAACGCTTCTGGACCAAGGCAAAATTCGAAAGCTGATGT	132
Db	432 CAAGCTTTGATGGCACCGCAACAAAGTCCCGCTGGGTGCAAAACGCAATCTCTTGCTGTTTC	491
Qy	133 GAGATGGAAGCGCACCGACAACAAATCTAAGTTGGGGGCAACGCCATCTGGGCGGTGC	192
Db	492 CATGGCTGTGCAAAAGGCTGCTGCTGATTTCCGAGGCTTCCCACTGCTTCCGCTACATCGG	551
Qy	193 CTTGGCTGTGCAAGGCTGGTGCAGCAGAGAAGGCGTCCCTCTACCGCACATCGC	252
Db	552 TGGACCAAAACG-----CACACGCTTCTTCAGTTTCAATGATGAACATCATCAACCG	602
Qy	253 CGACCTGGCGGGCAACCCCAAGATCATCTCCCGTTCCGCTTTTCAACGTCATCAACCG	312

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 09:39:18 ; Search time 38.5 Seconds
(without alignments)
824.048 Million cell updates/sec

Title: US-10-728-947-2
Perfect score: 2155
Sequence: 1 VAEIMHVAREILDSRGNT.....QLLDAGVYAGRSAPFRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/6C-COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/6D-COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	4	US-09-860-768-2
2	2151	99.8	425	4	US-09-860-768-4
3	1361.5	63.2	432	4	US-09-902-540-12060
4	1302.5	60.4	434	4	US-09-710-279-2222
5	1302.5	60.4	440	3	US-09-134-001C-5641
6	1282.5	59.5	444	4	US-09-252-991A-21020
7	1281.5	59.5	434	4	US-09-830-217-22
8	1281.5	59.5	434	4	US-10-278-946-22
9	1256.5	58.3	434	4	US-09-583-110-3817
10	1244	57.7	398	4	US-09-107-532A-3810
11	1232.5	57.2	444	4	US-09-328-352-7269
12	1226.5	56.9	420	3	US-08-847-065-19
13	1226.5	56.9	420	4	US-09-829-382-19
14	1226	56.9	459	4	US-09-489-039A-10061
15	1202.5	55.8	445	4	US-09-540-236-2354
16	1190.5	55.2	456	4	US-09-543-681A-6090
17	1167.5	54.2	434	4	US-09-438-185A-802
18	1131.5	52.5	407	4	US-09-107-433-4986
19	1109	51.1	443	4	US-09-949-016-8359
20	1099	51.0	434	4	US-09-949-016-6153
21	941.5	43.7	452	4	US-09-198-452A-853
22	902	41.9	311	3	US-08-961-083-184
23	902	41.9	311	4	US-09-536-784-184
24	654.5	30.4	301	4	US-08-248-796A-17053
25	504.5	23.4	177	1	US-08-446-920-11
26	458	21.3	148	4	US-09-248-796A-17054
27	265	12.3	113	4	US-09-513-999C-5420

28	224	10.4	84	4	US-09-198-452A-854	Sequence 854, Ap
29	214.5	10.0	70	3	US-08-847-065-17	Sequence 17, Appl
30	214.5	10.0	70	4	US-09-829-382-17	Sequence 17, Appl
31	214	9.9	80	4	US-09-513-999C-5931	Sequence 5931, Ap
32	206	9.6	74	4	US-09-513-999C-5841	Sequence 5841, Ap
33	181.5	8.4	70	4	US-09-829-382-16	Sequence 16, Appl
34	181.5	8.4	70	4	US-08-847-065-16	Sequence 16, Appl
35	156	7.2	48	3	US-08-715-034-2	Sequence 2, Appl
36	146	6.8	706	4	US-09-252-991A-30575	Sequence 30575, A
37	146	6.8	956	4	US-09-252-991A-17124	Sequence 17124, A
38	143	6.6	50	3	US-08-715-034-1	Sequence 1, Appl
39	119.5	5.5	718	4	US-09-252-991A-31865	Sequence 31865, A
40	119.5	5.5	1569	4	US-09-711-164-312	Sequence 312, Ap
41	117.5	5.5	523	4	US-09-252-991A-17991	Sequence 17991, A
42	112	5.2	472	4	US-09-134-000C-3893	Sequence 3893, Ap
43	112	5.2	1174	4	US-09-252-991A-29279	Sequence 29279, A
44	111.5	5.2	579	4	US-09-252-991A-29372	Sequence 29372, A
45	111	5.2	295	4	US-09-134-000C-4352	Sequence 4352, Ap

ALIGNMENTS

RESULT 1
US-09-860-768-2
; Sequence 2, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-2

Query Match	100.0%;	Score 2155;	DB 4;	Length 425;
Best Local Similarity	100.0%;	Pred. NO. 1.1e-189;		
Matches 425;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VAEIMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPVPGSGASTGVHAEHLRDCGDRYL	60	
Db	1	VAEIMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPVPGSGASTGVHAEHLRDCGDRYL	60	
Qy	61	KGVLKAVNNVEEIGDELADDDORLIDEAMTKLDGTANKSRKLGANALIGVSMVAKA	120	
Db	61	KGVLKAVNNVEEIGDELADDDORLIDEAMTKLDGTANKSRKLGANALIGVSMVAKA	120	
Qy	121	AADSAGLPLFRYIGGPNHVLVPMNMIINGGAHDSGVDPQFMIAPIGAETFEALRN	180	
Db	121	AADSAGLPLFRYIGGPNHVLVPMNMIINGGAHDSGVDPQFMIAPIGAETFEALRN	180	
Qy	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL	240	
Db	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVEAEKAGFTPGKDIALAL	240	
Qy	241	DVASSFFKDGTYHFEFGGQSAEMANVYAEVDVPIVSIEDPQDDHGEYTNLTATI	300	
Db	241	DVASSFFKDGTYHFEFGGQSAEMANVYAEVDVPIVSIEDPQDDHGEYTNLTATI	300	
Qy	301	GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVAHRAGYTSM	360	
Db	301	GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADVAHRAGYTSM	360	

QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
QY 421 PRFQG 425
DB 421 PRFQG 425

RESULT 2
US-09-860-768-4
; Sequence 4, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-4

Query Match 99.8%; Score 2151; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 2.6e-189;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
DB 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYL 60
QY 61 KGVLKAVENNEEIGDELAGEADDQRLIDEAMIKLDTGANKSRKLGANAILGVSMVAKA 120
DB 61 KGVLKAVENNEEIGDELAGEADDQRLIDEAMIKLDTGANKSRKLGANAILGVSMVAKA 120
QY 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETTFSEALRN 180
DB 121 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETTFSEALRN 180
QY 181 GAEVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGTTPGKDIALAL 240
DB 181 GAEVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGTTPGKDIALAL 240
QY 241 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTEDPLQEDDWEGYTNLTATI 300
DB 241 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTEDPLQEDDWEGYTNLTATI 300
QY 301 GDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRAGYTSM 360
DB 301 GDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRAGYTSM 360
QY 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
DB 361 SHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAP 420
QY 421 PRFQG 425
DB 421 PRFQG 425

RESULT 3
US-09-902-540-12060
; Sequence 12060, Application US/09902540

Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12060
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12060

Query Match 63.2%; Score 1361.5; DB 4; Length 432;
Best Local Similarity 64.0%; Pred. No. 1.1e-116;
Matches 272; Conservative 58; Mismatches 90; Indels 5; Gaps 2;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGG-DRYL 59
DB 1 MTEISQILAREVLDSRGNTVEAEVQLAGARGRAAVPSGASTGEHEAIELRDGDKHRYL 60
QY 60 KGVLKAVENNEEIGDELAGEADDQRLIDEAMIKLDTGANKSRKLGANAILGVSMVAK 119
DB 61 KGVLKAVENNEEIGDELAGEADDQRLIDEAMIKLDTGANKSRKLGANAILGVSMVAK 120
QY 120 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETTFSEALR 179
DB 120 AADSAGLPLFRYIGGPNHVLVPVPMNIIINGGAHADSGVDVQEPMTAPIGAETTFSEALR 180
QY 180 NGAEVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGTTPGKDIALA 239
DB 180 NGAEVYHALKSVIKKGLSTGLDGGFAPSFGSTREALDLIVEAIEKAGTTPGKDIALA 240
QY 240 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTEDPLQEDDWEGYTN 295
DB 240 DVASSEFFKQGYHFEQGQHSAAEMANYAELVDAYPIVSTEDPLQEDDWEGYTN 300
QY 296 LIATIGDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRA 355
DB 296 LIATIGDKVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMAHRA 360
QY 301 LTDALGSRMQLVGDDDLFVTNVERLGRGIETGTANSILVKVNOIGTLTETFDVDMAHRA 360
DB 301 LTDALGSRMQLVGDDDLFVTNVERLGRGIETGTANSILVKVNOIGTLTETFDVDMAHRA 360
QY 356 YTSMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYA 415
DB 356 YTSMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYA 420
QY 416 GRSAP 420
DB 421 GRSVF 425

RESULT 4
US-09-710-279-2222
; Sequence 2222, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348008
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2222
; LENGTH: 434
; TYPE: PRT

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2222

Query Match      60.4%; Score 1302.5; DB 4; Length 434;
Best Local Similarity 60.6%; Pred. No. 3e-111;
Matches 258; Conservative 65; Mismatches 94; Indels 9; Gaps 3;

Qy  4 IMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGK 62
Db  4 ITDYAREVILDSRGNTPTVEVEVLTESGAFGRALVPSGASTGEHEAVEHLRDGDKRYLKGK 63

Qy  63 VLKAVENNEIBDELAGLEAD--DQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
Db  64 VTKAVENNEMIAPEIVEGEFSLVDQVSDIKMWIQLDGTNKGKLGANAILGVSIARVA 123

Qy  121 AADSAGLPLFRYIGGPNNAHVLPVPMNIIINGGAHADSGVDVQEFMIAPIGAEFTSEALRN 180
Db  124 AADLLGQPLYKLGFGNGKQLPVPMMNIVNGGSHSDAPAFQEFMILPVGAESFKESLRW 183

Qy  181 GAETHYALKSVIKKGLSTGLDEGGFAPS VGSTREALLDLIVEATEKAGFTPGKDIALAL 240
Db  184 GAETHYALKSVIKKGLSTGLDEGGFAPS VGSTREALLDLIVEATEKAGFTPGKDIALAL 240

Qy  241 DVASSEFFKQGYTHF-----EGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYT 294
Db  244 DCASSEFYENGVDYTKFEGEGHAKRSAAEQVDYLEELIGKYPITIEDGMDENDWEGWK 303

Qy  295 NLTAIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMAHRA 354
Db  304 QLTDRIQDKVOLVGGDDLFTVNTTEILSKGIEGIGNSILIKVNOIGTTLTETFDIEMAOKA 363

Qy  355 GYTSMMSHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNQLLRIRIQLLDGAGVY 414
Db  364 GYTAVVSHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNQLLRIRIQLLDGAGVY 423

Qy  415 AGRSAF 420
Db  424 EGKSF 429

RESULT 5
US-09-134-001C-5641
; Sequence 5641, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5641
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5641

Query Match      60.4%; Score 1302.5; DB 3; Length 440;
Best Local Similarity 60.6%; Pred. No. 3.1e-111;
Matches 258; Conservative 65; Mismatches 94; Indels 9; Gaps 3;

Qy  4 IMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYLKGK 62
Db  10 ITDYAREVILDSRGNTPTVEVEVLTESGAFGRALVPSGASTGEHEAVEHLRDGDKRYLKGK 69

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2222

Query Match      59.5%; Score 1282.5; DB 4; Length 444;
Best Local Similarity 61.1%; Pred. No. 2.1e-109;
Matches 261; Conservative 59; Mismatches 100; Indels 7; Gaps 4;

Qy  1 VAEIMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYL 59
Db  16 MAKIVDIKREVLDSRGNTPTVEADVILDNLGIVGACAPSGASTGSRRALELRDGGKRYL 75

Qy  60 KGVLKAVENNEIBDELAGLEADQRLIDEAMIKLDGTANKSRIGANAILGVSMVAK 119
Db  76 KGVLKAVANINGPFRLLLGKDAADQKALDHAMELDTGTENKAKLGANAILAVSLAAK 135

Qy  120 AADSAGLPLFRYI-----GGPNAHVLPVPMNIIINGGAHADSGVDVQEFMIAPIGAEFTS 175
Db  136 AAAQAKGVPLFYAHTADLNGTPGQYMPVPMNIIINGGAHADNNVDIOEFMVQPVGAKNFA 195

Qy  176 EALRNGAEVHVALKSVIKKGLSTGLDEGGFAPS VGSTREALLDLIVEATEKAGFTPGKD 235
Db  196 EALRNGAEIIFHLKAVDKARGLNTAVGDEGFAPNLSSNEDALAAIAEAVKAGYKLGDD 255

Qy  236 IALALDVASSEFFKQGYTHFEG-QG-HSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGY 293

```

```

Qy  63 VLKAVENNEIBDELAGLEAD--DQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKA 120
Db  70 VTKAVENNEMIAPEIVEGEFSLVDQVSDIKMWIQLDGTNKGKLGANAILGVSIARVA 129

Qy  121 AADSAGLPLFRYIGGPNNAHVLPVPMNIIINGGAHADSGVDVQEFMIAPIGAEFTSEALRN 180
Db  130 AADLLGQPLYKLGFGNGKQLPVPMMNIVNGGSHSDAPAFQEFMILPVGAESFKESLRW 189

Qy  181 GAETHYALKSVIKKGLSTGLDEGGFAPS VGSTREALLDLIVEATEKAGFTPGKDIALAL 240
Db  190 GAETHYALKSVIKKGLSTGLDEGGFAPS VGSTREALLDLIVEATEKAGFTPGKDIALAL 249

Qy  241 DVASSEFFKQGYTHF-----EGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYT 294
Db  250 DCASSEFYENGVDYTKFEGEGHAKRSAAEQVDYLEELIGKYPITIEDGMDENDWEGWK 309

Qy  295 NLTAIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTTLTETFDVDMAHRA 354
Db  310 QLTDRIQDKVOLVGGDDLFTVNTTEILSKGIEGIGNSILIKVNOIGTTLTETFDIEMAOKA 369

Qy  355 GYTSMMSHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNQLLRIRIQLLDGAGVY 414
Db  370 GYTAVVSHRSGETEDTTIADLVALNCGQIKTGAPARSDRVAKYNQLLRIRIQLLDGAGVY 429

Qy  415 AGRSAF 420
Db  430 EGKSF 435

```

```

RESULT 6
US-09-252-991A-21020
; Sequence 21020, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21020
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21020

```

```

Query Match      59.5%; Score 1282.5; DB 4; Length 444;
Best Local Similarity 61.1%; Pred. No. 2.1e-109;
Matches 261; Conservative 59; Mismatches 100; Indels 7; Gaps 4;

Qy  1 VAEIMHVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGG-DRYL 59
Db  16 MAKIVDIKREVLDSRGNTPTVEADVILDNLGIVGACAPSGASTGSRRALELRDGGKRYL 75

Qy  60 KGVLKAVENNEIBDELAGLEADQRLIDEAMIKLDGTANKSRIGANAILGVSMVAK 119
Db  76 KGVLKAVANINGPFRLLLGKDAADQKALDHAMELDTGTENKAKLGANAILAVSLAAK 135

Qy  120 AADSAGLPLFRYI-----GGPNAHVLPVPMNIIINGGAHADSGVDVQEFMIAPIGAEFTS 175
Db  136 AAAQAKGVPLFYAHTADLNGTPGQYMPVPMNIIINGGAHADNNVDIOEFMVQPVGAKNFA 195

Qy  176 EALRNGAEVHVALKSVIKKGLSTGLDEGGFAPS VGSTREALLDLIVEATEKAGFTPGKD 235
Db  196 EALRNGAEIIFHLKAVDKARGLNTAVGDEGFAPNLSSNEDALAAIAEAVKAGYKLGDD 255

Qy  236 IALALDVASSEFFKQGYTHFEG-QG-HSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGY 293

```

```
Db      256  VTLALDCASSEFFKQKDYLEGCKVFDAAGFADYLAGLTQRYPIIISIEDGMDSDWAGW 315
Qy      294  TNLTATIGDKQIVGDDFFVFNTPRLKEGIKAANSILVKVQNIQIGTLTTFDAVDMHR 353
Db      316  KGLTDKIGAKVQVGGDLFVNTKILKEGIEKGIGNSILIKFNQIGSLTETLEAIQWAKA 375
Qy      354  AGYTSMMHSRSGETEDTTIADLAVALNCGOIKTGAPARSORVAKYNQLLRIEQLIGDAGV 413
Db      376  AGYTAIVSHRSGETEDTTIADLAVGTAAQGIKTSGLCSRDVSKYNQLLRIEQLGAKAP 435
Qy      414  YAGRSAP 420
Db      436  YRGRAEF 442

RESULT 7
US-09-830-217-22
; Sequence 22, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461PCT
; CURRENT APPLICATION NUMBER: US/09/830,217
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 22
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-830-217-22

Query Match      59.5%; Score 1281.5; DB 4; Length 434;
Best Local Similarity 59.1%; Pred. No. 2.6e-109;
Matches 253; Conservative 68; Mismatches 94; Indels 13; Gaps 4;

Qy      4  IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVGHEAHLRDGG-DRYLGKG 62
Db      4  ITDVAREVLDSRGNPTVEVEVLTSAGFGRALVPSGASTGEHEAHLRDGDKSRYLGKG 63
Qy      63  VLKAVENVNNEIIGDELAGLEAD----DQRLIDEAMIKLDGTANKSRILGANAILGVSMVA 118
Db      64  VTRAVENVNEIIAPEI--IEGEFSVLDOVSIDKMMIALDGTNPKGKLGANAILGVSTAVA 121
Qy      119  KAAADSAGLPFRYIGGPNNAHLVPMVMNIIINGGAHADSGVDVQEFMIAPIGAETFSAL 178
Db      122  RAAADLLGQPLYKLGFGNGKQLPVPMMNIVNGGSHSDAPIAQEFMILPVGATTFKESL 181
Qy      179  RNGAEVYHALKSVIKKEGLSTGLDEGGFAPSVGSTREALDLIVEALEKAGFTPGKDIAL 238
Db      182  RWGTEIFHNLSILSQRLGLETAVGDEGGFAPKEGTEDAVETIIQALEAAGYKPGEEVFL 241
Qy      239  ALDVASSEFFKQGTYPH-----EGGQHSAAEMANVYAEVLVDAYPIVISIEDPQEDDWE 292
Db      242  GFDCASEFFYGVYDYSKPEGEHGAKRTAAEQVDYLEQLVDKYPITIEDGMDENDWDG 301
Qy      293  YTNLTATIGDKVQIVGDDFFVFNTPRLKEGIKAANSILVKVQNIQIGTLTTFDAVDMH 352
Db      302  WKQLTERIGRVQLVGGDLFVTNTEILAKGIENGIGNSILIKVNIQIGTLTTFDAIEMAQ 361
Qy      353  RAGYTSMMHSRSGETEDTTIADLAVALNCGOIKTGAPARSORVAKYNQLLRIEQLIGDAG 412
Db      362  KAGYTAIVSHRSGETEDTTIADIAVATNAGQIKTGSLSRTDRIAKYNQLLRIEDELPE 421
Qy      413  YAGRSAP 420
Db      436  YRGRAEF 442

RESULT 9
US-09-583-110-3817
; Sequence 3817, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
```

```
Db      422  KYDGIKSF 429

RESULT 8
US-10-278-946-22
; Sequence 22, Application US/10278946
; Patent No. 6821754
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USD1
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-278-946-22

Query Match      59.5%; Score 1281.5; DB 4; Length 434;
Best Local Similarity 59.1%; Pred. No. 2.6e-109;
Matches 253; Conservative 68; Mismatches 94; Indels 13; Gaps 4;

Qy      4  IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVGHEAHLRDGG-DRYLGKG 62
Db      4  ITDVAREVLDSRGNPTVEVEVLTSAGFGRALVPSGASTGEHEAHLRDGDKSRYLGKG 63
Qy      63  VLKAVENVNNEIIGDELAGLEAD----DQRLIDEAMIKLDGTANKSRILGANAILGVSMVA 118
Db      64  VTRAVENVNEIIAPEI--IEGEFSVLDOVSIDKMMIALDGTNPKGKLGANAILGVSTAVA 121
Qy      119  KAAADSAGLPFRYIGGPNNAHLVPMVMNIIINGGAHADSGVDVQEFMIAPIGAETFSAL 178
Db      122  RAAADLLGQPLYKLGFGNGKQLPVPMMNIVNGGSHSDAPIAQEFMILPVGATTFKESL 181
Qy      179  RNGAEVYHALKSVIKKEGLSTGLDEGGFAPSVGSTREALDLIVEALEKAGFTPGKDIAL 238
Db      182  RWGTEIFHNLSILSQRLGLETAVGDEGGFAPKEGTEDAVETIIQALEAAGYKPGEEVFL 241
Qy      239  ALDVASSEFFKQGTYPH-----EGGQHSAAEMANVYAEVLVDAYPIVISIEDPQEDDWE 292
Db      242  GFDCASEFFYGVYDYSKPEGEHGAKRTAAEQVDYLEQLVDKYPITIEDGMDENDWDG 301
Qy      293  YTNLTATIGDKVQIVGDDFFVFNTPRLKEGIKAANSILVKVQNIQIGTLTTFDAVDMH 352
Db      302  WKQLTERIGRVQLVGGDLFVTNTEILAKGIENGIGNSILIKVNIQIGTLTTFDAIEMAQ 361
Qy      353  RAGYTSMMHSRSGETEDTTIADLAVALNCGOIKTGAPARSORVAKYNQLLRIEQLIGDAG 412
Db      362  KAGYTAIVSHRSGETEDTTIADIAVATNAGQIKTGSLSRTDRIAKYNQLLRIEDELPE 421
Qy      413  YAGRSAP 420
Db      422  KYDGIKSF 429

RESULT 9
US-09-583-110-3817
; Sequence 3817, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
```

RESULT 11
US-93-328-352-7269
; Sequence 7269, Application US/09328352
; Patent No. 5562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

; Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/829,382
 ; FILING DATE: 09-Apr-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/847,065
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-158 ..
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 420 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: S. pneumoniae
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Query Match 56.9%; Score 1226.5; DB 4; Length 420;
 Best Local Similarity 58.7%; Pred. No. 2.8e-104;
 Matches 244; Conservative 61; Mismatches 100; Indels 11; Gaps 4;
 Qy 15 SRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGG-DRYLGKGVKAVENNEE 73
 Db 1 SRGNPTLEVEVYTESGAFGRGMVPSGASTGHEAHLRDGKSRVGGIGTGKAVDNNVI 60
 Qy 74 IGDELAGEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAGSAGLPLFRYI 133
 Db 61 IADSIIGYDVRDQQAIDRAIMTALDGTNPKLGNAILGVSAVARAAADYLEIPLYSYL 120
 Qy 134 GGPNAHVLVPMNNIINGGAHSDGVVDQVFEMIAPIGAETPSEALRNGAEVYHAKSVIK 193
 Db 121 GGFNTKVLPTPMNNIINGGSHSDAPIAQFQEMILPVGAPTFKTLRYGAEIFHAKKILK 180
 Qy 194 EKGSTGLDGGFPAPSVGSTREALDLIVEAIEKAGTFPGKDIALDLVASESEFFK---- 249
 Db 181 SRGLETAGVDSGGAPREGTEGDGVETILAEAGVYVPGKDVFLGFCASSEFYDKERK 240
 Qy 250 --DGYTHFEG--GQHSAAENAVYAEVLVDYPIVSDIEDPLQEDDWEQYTNLTATIGDKV 304
 Db 241 VYDYT-KPEGGAAVRTSAEQIDYLELVNKYPIITIEDGNDENDWDGKALTRLGKKV 299
 Qy 305 QIVGDDRVNTPERLKEGIAKKAANSILVKNQIGTLTETFDVDMHACGYTSMMSHRS 364
 Db 300 QLVGDDFFVNTDYLARGIQGAANSILIKVKNQIGTLTETFEALMAKEAGYTAVSHRS 359
 Qy 365 GETEDTTIADLAVALNCCGQIKTGAPARSDRVAKYNQLRIEQLLDAGVYAGRSF 420
 Db 360 GETEDSTIADIATNAQGIKTSLSRTDRITAKYNQLRIEQLDGEVYAEVYGLKSF 415

RESULT 14
 US-09-489-039A-10061

; Sequence 10061, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 10061
 ; LENGTH: 459
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 ; US-09-489-039A-10061

Query Match 56.9%; Score 1226; DB 4; Length 459;
 Best Local Similarity 59.3%; Pred. No. 3.5e-104;
 Matches 252; Conservative 55; Mismatches 110; Indels 8; Gaps 3;
 Qy 1 VAEIMHVPAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGG-DRYL 59
 Db 28 MSKIVKVIIGREIIDSRGNPTVEAEVHLEGGFVGMAAAPSGASTGSRLELRDGDKSRFL 87
 Qy 60 GKVLKAVENNEEIGDELAGEADDQRLIDEAMIKLDGTANKSRILGANAILGVSMVAK 119
 Db 88 KGVTKAAVAVNGPIAQAILGKDAQAGIDKIMIDLDGTENKSNFGANAILAVFLANAK 147
 Qy 120 AAADSAGLPLFRYI----GGFNAHVLVPMNNIINGGAHSDGVVDQVFEMIAPIGAETFS 175
 Db 148 AAAASKGLPLYAHIAELNGTPGKYSMPVPMNNIINGGEHADNNVDIQEFMIQPVGAPTLK 207
 Qy 176 BALRNGAEVYHAKSVIKELSTGLDGGFAPSVGSTREALDLIVEAIEKAGFTPKD 235
 Db 208 EAVRMGSEVFHHLAKVLKSGMNTAVGDEGYAPNLGSAEALAVIAEAVKAAGVELGKD 267
 Qy 236 IALALDVASSEFFKDYTHFEGGQHSAA--AEMANVYAEVLVDYPIVSDIEDPLQEDDWEQ 292
 Db 268 ITLAWCARSSEFYKDKGVKVLAGEGNKAFTEEFTHFLEELTKQYPIVSDIEDGLDESDEW 327
 Qy 293 YTNLTATIGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKNQIGTLTETFDVDMHAK 352
 Db 328 FAYQTKVLGDKIQLVGGDLFVNTKILKEGIEKGIANSILIKFNQIGSLTETLAAIKNAK 387
 Qy 353 RAGYTSMSHRSRGTEDETTIADLAVALNCCGQIKTGAPARSDRVAKYNQLRIEQLLDGAG 412
 Db 388 DAGYTAIVISHRSGETEDATTIADLAVGTAAGQIKTGSMSRSDRVAKYNQLRIEELGEOA 447
 Qy 413 VYAGR 417
 Db 448 PFNGR 452

RESULT 15
 US-09-540-236-2354
 ; Sequence 2354, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540,236
 ; CURRENT FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 2354
 ; LENGTH: 445
 ; TYPE: PRT
 ; ORGANISM: M. catarrhalis
 ; US-09-540-236-2354

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:00:03 ; Search time 793.5 Seconds
(without alignments)
11772.340 Million cell updates/sec

Title: US-10-728-947-3

Perfect score: 1578

Sequence: 1 ggctgggatatggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1576.4	99.9	1578	4	Aaf61696 C. glut
2	1576.4	99.9	349980	5	Aah68527 C. glutami
3	1401.8	88.8	1405	10	Add13322 C. glutam
4	1394.8	88.4	1398	4	Aaf71395 Coryneb
5	1276.4	80.9	1278	8	Acc45311 Coryneb
6	1273.4	80.7	1275	5	Aah66050 C. glutami
7	1273.4	80.7	1299	8	Acc45312 Coryneb
8	1271.6	80.6	1278	8	Acc45314 Modified
9	1268.6	80.4	1299	8	Acc45313 Modified
10	907	57.5	1275	8	Acc29854 Prokaryot
11	678.2	43.0	1269	13	AdS56353 Bacterial
12	647	41.0	1287	8	ACA38079 Prokaryot
13	640.8	40.6	30159	4	Aas59564 Propionib
14	640.8	40.6	30159	8	ACF64493 Propionib
15	632.6	40.1	45190	6	Abx09144 Mycobacte
16	632.6	40.1	110000	4	Continuation (12 o
17	632.6	40.1	110000	4	Continuation (12 o
18	632.4	40.1	1317	8	ACA38642 Prokaryot
19	626.2	39.7	1290	8	ACA40436 Prokaryot
20	583.4	37.0	1344	8	ACA39628 Prokaryot

C	21	583	36.9	349980	6	ABQ81846
C	22	583	36.9	349980	6	ABQ81847
	23	580.4	36.8	1269	13	AD58570 Bacterial
	24	577	36.6	1269	13	AD58570 Bacterial
	25	575.6	36.5	1281	13	AD742160 Bacterial
	26	572.4	36.3	1275	13	AD5859832 Bacterial
	27	569	36.1	1260	13	AD5859832 Bacterial
	28	569	36.1	1269	13	AD5859832 Bacterial
	29	568.2	36.0	1278	13	AD743135 Bacterial
	30	567.4	36.0	1272	13	AD743135 Bacterial
	31	561.8	35.6	1272	13	AD585775 Bacterial
	32	558.6	35.4	1278	13	AD746979 Bacterial
	33	550.4	34.9	1290	13	AD744740 Bacterial
	34	542	34.3	1281	8	ACA23750 Prokaryot
	35	532	33.7	1335	11	ABD05845 Pseudomon
	36	531.4	33.7	1347	11	ABD05554 Pseudomon
C	37	530.2	33.6	1290	4	AAS54197 Pseudomon
	38	530.2	33.6	1290	8	ACA42456 Prokaryot
	39	529.2	33.5	1281	13	AD746255 Bacterial
	40	528	33.5	1287	13	AD5856507 Bacterial
	41	522.8	33.1	1278	13	AD549945 Bacterial
	42	522.8	33.1	1281	8	ACA25228 Prokaryot
	43	522.8	33.1	1281	13	AD585460 Bacterial
	44	515.8	32.7	1690	12	ADO59752 B. subtil
	45	513.6	32.5	1404	8	ACA45728 Prokaryot

ALIGNMENTS

RESULT 1
AAF61696
ID AAF61696 standard; DNA; 1578 BP.

AC AAF61696;

DT 12-JUL-2001 (first entry)

DE C. glutamicum enolase encoding DNA.

XX Enolase; fermentation; L-amino acid; L-lysine; coryneform; eno gene;
KW medicine; animal feed supplement; ds.

OS Corynebacterium glutamicum.

FH Key Location/Qualifiers
CDS 151..1428

FT /*tag= a
FT /product= "Enolase"

XX EP1090998-AL.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121158.

XX 05-OCT-1999; 99DE-01047791.

XX (DEGS) DEGUSSA-HUELS AG.

XX Moeckel B, Pfeifferle W, Hermann T, Puehler A, Kalinowski J;
XX Bathe B;

XX WPI; 2001-292928/31.

XX P-PSDB; AAB70882.

XX New enolase gene from coryneform bacteria, used to prepare transformants
PT with increased synthesis of amino acids, particularly lysine.

XX Claim 4; Page 12-14; 25pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
CC coryneform bacteria which is used in a method for fermentative production

CC of L-amino acids, especially L-lysine, by fermenting a lysine-producing
 CC coryneform in which the eno gene has been amplified, and isolating amino
 CC acids that have accumulated in the medium or cells. (1), which encodes an
 CC enolase, is used to transform coryneforms for production of L-amino
 CC acids, specifically lysine which is used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Overexpression of (1) improves production of
 CC amino acids, especially of L-lysine. This sequence encodes the enolase
 CC protein described in the method of the invention
 XX
 SQ Sequence 1578 BP; 351 A; 461 C; 424 G; 342 T; 0 U; 0 Other;

Query Match 99.98; Score 1576.4; DB 4; Length 1578;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTTGCCCTCATCGAAACAAGA 60
 DB 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAATTTCAACTGATTTGCCCTCATCGAAACAAGA 60
 QY 61 TTGCTGCAACAATTTGGGTAGTAGCTGATTTGAAGACATTTTGTATGATCATCGTGAATATTCTAG 120
 DB 61 TTGCTGCAACAATTTGGGTAGTAGCTGATTTGAAGACATTTTGTATGATCATCGTGAATATTCTAG 120
 QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
 DB 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
 QY 181 GAAATTCCTGACTCCCGCGGTAAACCAACCGTTCAGGCGAGAGGTTTTCCTGGATGACGGT 240
 DB 181 GAAATTCCTGACTCCCGCGGTAAACCAACCGTTCAGGCGAGAGGTTTTCCTGGATGACGGT 240
 QY 241 TCCACAGGTTCGAGGTGTTTCATTCGGCGCATTCACCGCGGTTCACGAGGCTCATGAG 300
 DB 241 TCCACAGGTTCGAGGTGTTTCATTCGGCGCATTCACCGCGGTTCACGAGGCTCATGAG 300
 QY 301 CTGGCTCAGGTGGCGATCGCTACTGGCAAGGGCGTTTGAAGGAGTTGAAGAGTC 360
 DB 301 CTGGCTCAGGTGGCGATCGCTACTGGCAAGGGCGTTTGAAGGAGTTGAAGAGTC 360
 QY 361 AACGAAGAAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420
 DB 361 AACGAAGAAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGAC 420
 QY 421 GAAGCAATGATCAAGCTTGATGGACCGCCAAACGATCCCGCTGGGTGCAAGCGCAATC 480
 DB 421 GAAGCAATGATCAAGCTTGATGGACCGCCAAACGATCCCGCTGGGTGCAAGCGCAATC 480
 QY 481 CTTGGTCTTCCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTC 540
 DB 481 CTTGGTCTTCCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTC 540
 QY 541 CGCTACATCGGTGACCAAGCGACAGCTTCTCCAGTTTCCAAATGATGAACATCATCAAC 600
 DB 541 CGCTACATCGGTGACCAAGCGACAGCTTCTCCAGTTTCCAAATGATGAACATCATCAAC 600
 QY 601 GGTGGCGTCAAGCTGATCGCTGGGTGATGAGTTTCAAGAAATTCATGATCGCTCCAAATCGGT 660
 DB 601 GGTGGCGTCAAGCTGATCGCTGGGTGATGAGTTTCAAGAAATTCATGATCGCTCCAAATCGGT 660
 QY 661 GCAGAGACCTTCTCTGAGGCTCTCGCAAGCGGCGGAGGCTTACACGCACTGAAGTCC 720
 DB 661 GCAGAGACCTTCTCTGAGGCTCTCGCAAGCGGCGGAGGCTTACACGCACTGAAGTCC 720
 QY 721 GTCATCAAGGAAAGGGCTCTGCAACCGGATTCGGCGATGAGGGCGCTTCGCTCCCTTCC 780
 DB 721 GTCATCAAGGAAAGGGCTCTGCAACCGGATTCGGCGATGAGGGCGCTTCGCTCCCTTCC 780
 QY 781 GTGCGGCTCCACCGGTGAGGCTCTTGAACCTTATCGTTAAGGCAATCGAAGGCTGGCTTC 840
 DB 781 GTGCGGCTCCACCGGTGAGGCTCTTGAACCTTATCGTTAAGGCAATCGAAGGCTGGCTTC 840
 QY 841 ACCCCAGGCAAGGACATCGCTCTTGCTCTGAGCGTTGCTTCTCTCTGAGTTCTTCAAGGAC 900

DB 841 ACCCAGGCAAGGACATCGCTCTTGCTTGGAGTTGCTTCTCTGAGTTCTTCAAGGAC 900
 QY 901 GGCACCTACCACTTCAAGGTGGCCAGCACTCCGAGCTGAGATGCAAAACGTTTACGCT 960
 DB 901 GGCACCTACCACTTCAAGGTGGCCAGCACTCCGAGCTGAGATGCAAAACGTTTACGCT 960
 QY 961 GAGCTGTTGACCGCTACCAATCGTCTCATCGAGGACCCACTGCGAGGAGTACTGG 1020
 DB 961 GAGCTGTTGACCGCTACCAATCGTCTCATCGAGGACCCACTGCGAGGAGTACTGG 1020
 QY 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCAAGGTTTCAGATCGTTGCGGACGAC 1080
 DB 1021 GAGGTTTACACCAACCTCACCGCAACCATCGGCAAGGTTTCAGATCGTTGCGGACGAC 1080
 QY 1081 TTCTTCGTCAACCAACCTGAGCGCTGGAAGGAGGGCATCGCTAAGAGGCTGCAACTCC 1140
 DB 1081 TTCTTCGTCAACCAACCTGAGCGCTGGAAGGAGGGCATCGCTAAGAGGCTGCAACTCC 1140
 QY 1141 ATCTCTGTTAAGGTGAACCAAGTCCGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
 DB 1141 ATCTCTGTTAAGGTGAACCAAGTCCGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
 QY 1201 GCTCACGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAAGGAGGACACC 1260
 DB 1201 GCTCACGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCGTGAAGGAGGACACC 1260
 QY 1261 ACCATTGCTGACCTCGAGTTGCACTCAATGTGGCGAGATCAAGACTGGTGTCCAGCA 1320
 DB 1261 ACCATTGCTGACCTCGAGTTGCACTCAATGTGGCGAGATCAAGACTGGTGTCCAGCA 1320
 QY 1321 CGTTCCGACCGGTTCGCAAGTACACCAAGTTCGCGATCGAGCAGCTGTTGGCGAC 1380
 DB 1321 CGTTCCGACCGGTTCGCAAGTACACCAAGTTCGCGATCGAGCAGCTGTTGGCGAC 1380
 QY 1381 GCGGCGTCTACGAGGTGCGCAGCATTCGCCACCGCTTTCAGGGCTTAAATAAAGCGCTT 1440
 DB 1381 GCGGCGTCTACGAGGTGCGCAGCATTCGCCACCGCTTTCAGGGCTTAAATAAAGCGCTT 1440
 QY 1441 TTCGACGCGCGTAACTCAAGTTTCCGCGGCTGCTGCTTACTACTGTTACTGTTGT 1500
 DB 1441 TTCGACGCGCGTAACTCAAGTTTCCGCGGCTGCTGCTTACTACTGTTACTGTTGT 1500
 QY 1501 GACTATGATCGAGGATTTATGGCAAGCAAGCAAGAACTCATAAAGGCTTGTTCCTGCT 1560
 DB 1501 GACTATGATCGAGGATTTATGGCAAGCAAGCAAGAACTCATAAAGGCTTGTTCCTGCT 1560
 QY 1561 CAAGCAGGGAACCGTCTT 1578
 DB 1561 CAAGCAGGGAACCGTCTT 1578
 RESULT 2
 AAH68527
 ID AAH68527 standard; DNA; 349980 BP.
 XX AAH68527;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 XX C glutamicum coding sequence fragment SEQ ID NO: 7062.
 DE
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EF1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 XX 18-DEC-2000; 2000EP-00127688.
 PF
 XX

PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI; 2001-376931/40.
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX Disclosure; SEQ ID NO 7062; 246pp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 349980 BP; 79725 A; 90426 C; 98918 G; 80911 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1576.4; DB 5; Length 349980;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GGCTGGGGATGATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
 Db 134799 GGCTGGGGATGATGGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 134858
 Qy 61 TTCTGTCAACAATTTGGGTGTAGAGCGTATTGAAGACATTTGATCAGTGAATAATTTCTAG 120
 Db 134859 TTCTGTCAACAATTTGGGTGTAGAGCGTATTGAAGACATTTGATCAGTGAATAATTTCTAG 134918
 Qy 121 TTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATCGATTTTCGCTCGC 180
 Db 134919 TTAGTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATCGATTTTCGCTCGC 134978
 Qy 181 GAAATTTCTCGACTCCCGCGTAAACCAACCGTCCGAGGAGGATTTTCCTGGATGACGGT 240
 Db 134979 GAAATTTCTCGACTCCCGCGTAAACCAACCGTCCGAGGAGGATTTTCCTGGATGACGGT 135038
 Qy 241 TCCACCGGTGTCAGGTGTTCCATCCGCGGCATCCACCGCGTCCACGAGGCTCATGAG 300
 Db 135039 TCCACCGGTGTCAGGTGTTCCATCCGCGGCATCCACCGCGTCCACGAGGCTCATGAG 135098
 Qy 301 CTGGGTGACGTTGGCATCGTACTCTGGGCAAGGCGTTTGAAGGACGATTTGAACAGTC 360
 Db 135099 CTGGGTGACGTTGGCATCGTACTCTGGGCAAGGCGTTTGAAGGACGATTTGAACAGTC 135158
 Qy 361 AACGAAGAAATCGCGGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCGCTCATCGAC 420
 Db 135159 AACGAAGAAATCGCGGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCGCTCATCGAC 135218
 Qy 421 GAAGCAATGATCAAGCTTTGATGGCACCGCCAAAGTCCCGCTGGGTGCAAAACGCAATC 480
 Db 135219 GAAGCAATGATCAAGCTTTGATGGCACCGCCAAAGTCCCGCTGGGTGCAAAACGCAATC 135278
 Qy 481 CTTGGTGTTCATGGCTGTTGCAAAAGTCTGCTGATTCGCGAGGCTCCCACTGTTC 540
 Db 135279 CTTGGTGTTCATGGCTGTTGCAAAAGTCTGCTGATTCGCGAGGCTCCCACTGTTC 135338
 Qy 541 CGCTACATCGGTGACCAAGCGACACGTTCTTCCAGTTCCAAATGATCAATCATCAAC 600

Db	135339	CGCTACATCGGTGACCAAAACGACACGTTCTTCCAGTTCCAAATGATCAATCATCAAC	135398
Qy	601	GGTGGCGCTCAGCTGACTCGGTGTTGAGTTCAGGAATTCATGATCGCTCCATCGGT	660
Db	135399	GGTGGCGCTCAGCTGACTCGGTGTTGAGTTCAGGAATTCATGATCGCTCCATCGGT	135458
Qy	661	GCAGAGACCTTTCTGAGGGCTCTCCGCAACGGCGCGGAGGTCTACCAAGTCC	720
Db	135459	GCAGAGACCTTTCTGAGGGCTCTCCGCAACGGCGCGGAGGTCTACCAAGTCC	135518
Qy	721	GTCAATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCTTCC	780
Db	135519	GTCAATCAAGGAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTTCGCTCTTCC	135578
Qy	781	GTGGGCTCCACCGCTGAGGGCTCTGACCTTATCGTTAAGCAATCGAAGCGCTGGCTTC	840
Db	135579	GTGGGCTCCACCGCTGAGGGCTCTGACCTTATCGTTAAGCAATCGAAGCGCTGGCTTC	135638
Qy	841	ACCCAGGCAAGGACATCGCTCTTGTCTGTGACCTTGTCTCTCTGAGTTCTTCAAGGAC	900
Db	135639	ACCCAGGCAAGGACATCGCTCTTGTCTGTGACCTTGTCTCTCTGAGTTCTTCAAGGAC	135698
Qy	901	GGCACTTACCACTTCGAAGGTGGCAGCATCTCCGAGCTGAGATGGCAAAAGTTCAGCT	960
Db	135699	GGCACTTACCACTTCGAAGGTGGCAGCATCTCCGAGCTGAGATGGCAAAAGTTCAGCT	135758
Qy	961	GAGCTCGTTGACGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG	1020
Db	135759	GAGCTCGTTGACGGTACCAATCGTCTCCATCGAGGACCCACTCGAGGAAGATGACTGG	135818
Qy	1021	GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACAC	1080
Db	135819	GAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACAC	135878
Qy	1081	TTCTTCTGTCACCAACCTCGAGCGCTGAAAGAGGGCATCGCTAAGAGAGGCTGCCAACTCC	1140
Db	135879	TTCTTCTGTCACCAACCTCGAGCGCTGAAAGAGGGCATCGCTAAGAGAGGCTGCCAACTCC	135938
Qy	1141	ATCTTGTGTTAAGGTGAACAGATCGGTACCTCACCGAGACTTCGAGCGCTGTGACATG	1200
Db	135939	ATCTTGTGTTAAGGTGAACAGATCGGTACCTCACCGAGACTTCGAGCGCTGTGACATG	135998
Qy	1201	GCTCACCGCGAGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACAC	1260
Db	135999	GCTCACCGCGAGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACCGAGGACAC	136058
Qy	1261	ACCATTTGCTGACCTTCGAGTTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Db	136059	ACCATTTGCTGACCTTCGAGTTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	136118
Qy	1321	CGTTCGACCGTGTGCAAAAGTACAACAGCTTCTCCGATCGAGAGCTGCTTGGCGAC	1380
Db	136119	CGTTCGACCGTGTGCAAAAGTACAACAGCTTCTCCGATCGAGAGCTGCTTGGCGAC	136178
Qy	1381	GCCGGCGCTTACGAGGTTCGAGCGCATTCACCGCTTTCAGGGCTAAATAAAGCGCTT	1440
Db	136179	GCCGGCGCTTACGAGGTTCGAGCGCATTCACCGCTTTCAGGGCTAAATAAAGCGCTT	136238
Qy	1441	TTTGACCGCCCGGTAAACCTCAAGGTTTGGCGGCGCTGTTTGGCTTACTACTGTTACTGTGT	1500
Db	136239	TTTGACCGCCCGGTAAACCTCAAGGTTTGGCGGCGCTGTTTGGCTTACTACTGTTACTGTGT	136298
Qy	1501	GACTATGATCGAGATTATGGCAAGCAGAGAAAACCTCATAAAGGCTTGTCTGTCT	1560
Db	136299	GACTATGATCGAGATTATGGCAAGCAGAGAAAACCTCATAAAGGCTTGTCTGTCT	136358
Qy	1561	CAAGCAGGAAACGTGCTT 1578	
Db	136359	CAAGCAGGAAACGTGCTT 136376	

RESULT 3

[illegible]

QY 1191 TGTGACATGGCTACCCGCGAGGCTACACCTCATGATGTCCACACGTTCCGGTGAGAC 1250
 Db 1141 TGTGACATGGCTACCCGCGAGGCTACACCTCATGATGTCCACACGTTCCGGTGAGAC 1200
 QY 1251 CGAGGACACCACTTGTGACCTGCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCACTTGTGACCTGCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
 QY 1311 TGCTCCAGCAGCTTCCGACCGTGTGCGAAAGTCAACACGAGCTTCTCCGCTCAGGACGCT 1370
 Db 1261 TGCTCCAGCAGCTTCCGACCGTGTGCGAAAGTCAACACGAGCTTCTCCGCTCAGGACGCT 1320
 QY 1371 GCTTGGCAGCGCGGCTTACGAGGTCGAGGCAATCCACGCTTTCAGGCTAAAT 1430
 Db 1321 GCTTGGCAGCGCGGCTTACGAGGTCGAGGCAATCCACGCTTTCAGGCTAAAT 1380
 QY 1431 AAAAGCGCTTTTCGACGCCCGGTAA 1455
 Db 1381 AAAAGCGCTTTTCGACGCCCGGTAA 1405

RESULT 4

AAAF71395

ID AAF71395 standard; DNA; 1398 BP.

XX AAF71395;

AC AAF71395;

DT 30-APR-2001 (first entry)

XX Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:71.
 DE Corynebacterium glutamicum; carbon metabolism and energy production;
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
 XX

OS Corynebacterium glutamicum.

XX

FN WO200100844-A2.

XX

XX

PD 04-JAN-2001.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042123.
 PR 03-SEP-1999; 99DE-01042125.
 XX (BADI) BASF AG.
 FA Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-061975/07.
 PI P-PSDB; AAB79278.
 DR
 DR
 DR
 XX

New isolated Corynebacterium glutamicum nucleic acid encoding a sugar metabolism and oxidative phosphorylation protein for production or modulation of production of fine chemicals e.g. amino acids, carbohydrates or enzymes.

Claim 3; Page 242-244; 1246pp; English.

XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243 to AAB 79633 which are involved in carbon metabolism and energy production. The C. glutamicum SMP gene can be used in vectors (II) for expression in host cells and production or modulation of production of fine chemicals, such as, an organic acid, a proteinogenic or nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins (III) encoded by them are used for diagnosing the presence or activity of Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells containing them are used to map genomes of organisms related to C. glutamicum, identify and localise C. glutamicum sequences of interest, in evolutionary studies, in determining SMP protein regions required for function, in modulating SMP protein activity, in modulating the metabolism of sugars, and in modulating high-energy molecule production in a cell (i.e. ATP, NADPH)

SQ Sequence 1398 BP; 308 A; 423 C; 376 G; 291 T; 0 U; 0 Other;

Query Match 88.4%; Score 1394.8; DB 4; Length 1398;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 CGAACAAGATTGTCGCAACAATTGGGTGTAGAGTGTGAAGACATTTGATCACGTGA 110
 Db 1 CGAACAAGATTGTCGCAACAATTGGGTGTAGAGTGTGAAGACATTTGATCACGTGA 60
 QY 111 ATAAATCTAGTTCCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGT 170
 Db 61 ATAAATCTAGTTCCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGT 120
 QY 171 ATTGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCCACCGTCGAGGACAGGTTTTTCT 230
 Db 121 ATTGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCCACCGTCGAGGACAGGTTTTTCT 180
 QY 231 GGATGACGGTTCCGACGGTGTCCAGGTGTTCATCCGCGCATCCACCGCGTCCACGA 290
 Db 181 GGATGACGGTTCCGACGGTGTTCATCCGCGCATCCACCGCGTCCACGA 240
 QY 291 GGCTCATGAGCTCGCTGACGGTGGCGATCGTACCTGGGCAAGGGCGTTTTGAAGGCGAGT 350
 Db 241 GGCTCATGAGCTCGCTGACGGTGGCGATCGTACCTGGGCAAGGGCGTTTTGAAGGCGAGT 300
 QY 351 TGAACAACGTCACGAAGAATTCGCGACGAGCTCGCTGGCTTAGAGGCTCAGCATCAGCG 410
 Db 301 TGAACAACGTCACGAAGAATTCGCGACGAGCTCGCTGGCTTAGAGGCTCAGCATCAGCG 360
 QY 411 CCTCATCGAGCAAGCAATGATCAAGCTTGTATGGCAGCGCCAAAGTCCCGCTGGGTGC 470
 Db 361 CCTCATCGAGCAAGCAATGATCAAGCTTGTATGGCAGCGCCAAAGTCCCGCTGGGTGC 420
 QY 471 AAACGCAATCCTTGGTGTTCATGGCTGTGTGAAAGGCTGCTGCTGATTCGCGAGGCT 530

Db 421 AAAGCAGTCTTGGTGTTCATGGCTGTGCAAGGCTGCTGATTCGACGGCT 480
 Qy 531 CCACATGTTCCGTTACATCGGTGGAGCAAAACGACACAGTTCTTCCAGTTCCAAATGATGAA 590
 Db 481 CCACATGTTCCGTTACATCGGTGGAGCAAAACGACACAGTTCTTCCAGTTCCAAATGATGAA 540
 Qy 591 CATCATCAAGCGTGGCGCTCACGCTGACTCCGTTGAGTTGAGTTTCAGGAATTCATGATCGC 650
 Db 541 CATCATCAAGCGTGGCGCTCACGCTGACTCCGTTGAGTTTCAGGAATTCATGATCGC 600
 Qy 651 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGC 710
 Db 601 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGC 660
 Qy 711 ACTGAAGTCCGTCATCAAGAAAGGGCTGTGTCCACCGGACTTGGCGATGAGGGGCGGCTT 770
 Db 661 ACTGAAGTCCGTCATCAAGAAAGGGCTGTGTCCACCGGACTTGGCGATGAGGGGCGGCTT 720
 Qy 771 CGCTCCTTCCGTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
 Db 721 CGCTCCTTCCGTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780
 Qy 831 GGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGCTCTGGAAGTTCTTCTCTGAGTT 890
 Db 781 GGCTGGCTTCAACCCAGGCAAGACATCGCTCTTGCTCTGGAAGTTCTTCTCTGAGTT 840
 Qy 891 CTTCAAGGACGGACCTTACCACTTCGAAAGTGGCCAGCATCTCCGAGCTGAGATGGCAAA 950
 Db 841 CTTCAAGGACGGACCTTACCACTTCGAAAGTGGCCAGCATCTCCGAGCTGAGATGGCAAA 900
 Qy 951 CGTTAGCTGAGTCTGCTGACCGGTACCAACCTTCCGCAACCACTCGAGGACCACTTCGAGGA 1010
 Db 901 CGTTAGCTGAGTCTGCTGACCGGTACCAACCTTCCGCAACCACTCGAGGACCACTTCGAGGA 960
 Qy 1011 AGATGACTGGAGGGTTACACCAACCTTCCGCAACCACTCGAGGACCACTTCGAGTCTGT 1070
 Db 961 AGATGACTGGAGGGTTACACCAACCTTCCGCAACCACTCGAGGACCACTTCGAGTCTGT 1020
 Qy 1071 TGGCAGCAGCTTCTGCTACCAACCTTGAGCGCTGAAAGGAGGCAATCGCTAAGAGGC 1130
 Db 1021 TGGCAGCAGCTTCTGCTACCAACCTTGAGCGCTGAAAGGAGGCAATCGCTAAGAGGC 1080
 Qy 1131 TGGCACTCCATCTGCTGAGTGAACCAAGTACCGTACCGTACCGAGACCTTCGACGC 1190
 Db 1081 TGGCACTCCATCTGCTGAGTGAACCAAGTACCGTACCGTACCGAGACCTTCGACGC 1140
 Qy 1191 TGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCCGTTGAGAC 1250
 Db 1141 TGTGACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGTTCCGTTGAGAC 1200
 Qy 1251 CGAGGACACCACTTCTGACCTCGAGTTGCACTCAATGTGGCCAGATCAAGACTGG 1310
 Db 1201 CGAGGACACCACTTCTGACCTCGAGTTGCACTCAATGTGGCCAGATCAAGACTGG 1260
 Qy 1311 TGTCCAGCAGCTTCGACCGGTGTCGCAAGTACCAACGAGTCTCCGCAATCGAGCAGCT 1370
 Db 1261 TGTCCAGCAGCTTCGACCGGTGTCGCAAGTACCAACGAGTCTCCGCAATCGAGCAGCT 1320
 Qy 1371 GCTTGGCGAGCCCGGCTCTACGAGGTTCGAGCGGCAATTCACCGCTTTTCAGGCTTAAT 1430
 Db 1321 GCTTGGCGAGCCCGGCTCTACGAGGTTCGAGCGGCAATTCACCGCTTTTCAGGCTTAAT 1380
 Qy 1431 AAAAGCGCTTTTCGACGC 1448
 Db 1381 AAAAGCGCTTTTCGACGC 1398

RESULT 5
 ACC45311
 ID ACC45311 standard; DNA; 1278 BP.
 XX
 AC ACC45311;

XX DT 17-JUN-2003 (first entry)
 XX DE Corynebacterium glutamicum eno encoding DNA SEQ ID NO:1.
 XX KW Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
 KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
 KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
 KW pharmaceutical; food; animal feeding; eno; gene; ds.
 XX OS Corynebacterium glutamicum.
 XX Key Location/Qualifiers
 CDS 1..1278
 FT /*tag= a
 FT /product= "eno protein"
 XX PN WO2003023016-A2.
 XX PD 20-MAR-2003.
 XX PF 11-SEP-2002; 2002WO-EP010174.
 XX PR 13-SEP-2001; 2001DE-01045043.
 XX PA (DEGS) DEGUSSA AG.
 XX PI Farwick M, Hermann T;
 XX DR WPI; 2003-354534/33.
 XX DR P-PSDB; ABP97001.
 XX PT Microorganism useful for producing e.g. fine chemicals, has permanently
 PT altered phosphorylatability protein, such that biosynthesis of fine
 PT chemical synthesized by microorganism is increased compared to wild-type.
 XX Example 3; Page 78-80; 120pp; English.
 CC The present invention describes a microorganism (I), in which the
 CC phosphorylatability of at least one protein has been permanently altered
 CC such that the biosynthesis of at least one fine chemical synthesised by
 CC the microorganism is increased compared to the wild type. Also described:
 CC (1) use of a DNA (II) sequence coding for a protein which contains a
 CC phosphorylation site, where the sequence contains such a mutation that
 CC the protein is changed in its phosphorylatability for the production of
 CC (1), or for the production of fine chemicals; and (2) a method for
 CC producing fine chemicals or metabolites comprising using (I). (I) is
 CC useful for producing fine chemicals or metabolites, such as amino acids,
 CC vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
 CC and vitamins produced using (I) can be used in human medicine, in the
 CC pharmaceutical industry, food industry and in animal feeding. (I)
 CC produces larger amount of desired fine chemical or a metabolite than the
 CC wild type. The present sequence encodes wild type eno from
 CC Corynebacterium glutamicum, which is used in an example from the present
 CC invention
 XX SQ Sequence 1278 BP; 271 A; 399 C; 347 G; 261 T; 0 U; 0 Other;
 Query Match 80.9%; Score 1276.4; DB 8; Length 1278;
 Best Local Similarity 99.9%; Pred. No. 2.8e-313;
 Matches 1277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 151 GTGGCTGAATCATGACGTTATTCGTCGCGAAATTCGACTCCCGCGGTACCCCAACC 210
 Db 1 GTGGCTGAATCATGACGTTATTCGTCGCGAAATTCGACTCCCGCGGTACCCCAACC 60
 Qy 211 GTCGAGCGCAGAGGTTTCTTGGATGACGGTTCACAGGTTCCAGAGTGTTCATCCGGC 270
 Db 61 GTCGAGCGCAGAGGTTTCTTGGATGACGGTTCACAGGTTCCAGAGTGTTCATCCGGC 120
 Qy 271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTCGGTGACCGGTGCGGATCGCTACCTGGGC 330
 Db 121 GCATCCACCGCGCTCCACGAGGCTCATGAGCTCGGTGACCGGTGCGGATCGCTACCTGGGC 180

```

Qy 331 AAGGCGCTTTTGAAGGCGAGTTGAAAACTCAACGAGAAATCGGCGACGAGCTCGCTGGC 390
Db 181 AAGGCGCTTTTGAAGGCGAGTTGAAAACTCAACGAGAAATCGGCGACGAGCTCGCTGGC 240

Qy 391 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGTATGCGACCGCC 450
Db 241 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGTATGCGACCGCC 300

Qy 451 AACAAATCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGAAAGCT 510
Db 301 AACAAATCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGAAAGCT 360

Qy 511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGTACATCGGTGACCAACGACACGTT 570
Db 361 GCTGCTGATTCGCGAGCGCTCCCACTGTTCCGTACATCGGTGACCAACGACACGTT 420

Qy 571 CTTCCAGTTCATATGATGAACATCATCAACGGTGGCGCTCACGCTGATCCGGTGTGAC 630
Db 421 CTTCCAGTTCATATGATGAACATCATCAACGGTGGCGCTCACGCTGATCCGGTGTGAC 480

Qy 631 GTTCAGGAATTCATGATCGCTCCCAATCGGTGCGAGACCTCTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGGAATTCATGATCGCTCCCAATCGGTGCGAGACCTCTCTGAGGCTCTCCGCAAC 540

Qy 691 GGGCGGAGGTCTACCAAGCACTGAAGTCCGTCTCAAGGAAAGGGCCTGTCCACCGGA 750
Db 541 GGGCGGAGGTCTACCAAGCACTGAAGTCCGTCTCAAGGAAAGGGCCTGTCCACCGGA 600

Qy 751 CTTGGCGATGAGGCGGCTTTCGCTTCCGTTCGGTCCACCGGTGAGGCTCTTTGACCTT 810
Db 601 CTTGGCGATGAGGCGGCTTTCGCTTCCGTTCGGTCCACCGGTGAGGCTCTTTGACCTT 660

Qy 811 ATCGTTAAGCAATCGAGAGGCTGGCTTCAACCGGCGAAGGACATCGCTTTCGCTG 870
Db 661 ATCGTTAAGCAATCGAGAGGCTGGCTTCAACCGGCGAAGGACATCGCTTTCGCTG 720

Qy 871 GAGCTTCTCTCTGAGTCTTCAAGGAGGCGACCTTACCACTTCGAAGGTGGCGACAC 930
Db 721 GAGCTTCTCTCTGAGTCTTCAAGGAGGCGACCTTACCACTTCGAAGGTGGCGACAC 780

Qy 931 TCCGACGCTGAGATGGCAAGCTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 990
Db 781 TCCGACGCTGAGATGGCAAGCTTTACGCTGAGCTCGTTGACGCGTACCAATCGTCTCC 840

Qy 991 ATCGAGGACCACTGCGAGGAGATGACTGGAGGGTTACCAACCTTACCGCAACCATC 1050
Db 841 ATCGAGGACCACTGCGAGGAGATGACTGGAGGGTTACCAACCTTACCGCAACCATC 900

Qy 1051 GGCACAAAGTTTCAGATCGTTGGCGAGGACTTCTTCGTTCACCAACCTGAGCGCTGAAG 1110
Db 901 GGCACAAAGTTTCAGATCGTTGGCGAGGACTTCTTCGTTCACCAACCTGAGCGCTGAAG 960

Qy 1111 GAGGCACTCGCTAAGAGGCTGCCAATCCATCTCGTTAAGGTGAACAGATCGGTACC 1170
Db 961 GAGGCACTCGCTAAGAGGCTGCCAATCCATCTCGTTAAGGTGAACAGATCGGTACC 1020

Qy 1171 CTACCGAGACCTTTCGACGCTGTCGATGCTGCTCACCAGGAGGCTACCTTCATGATG 1230
Db 1021 CTACCGAGACCTTTCGACGCTGTCGATGCTGCTCACCAGGAGGCTACCTTCATGATG 1080

Qy 1231 TCCACCGCTTCGGTGAGACCGAGGACCAACCAATTCGCTCAGCTTCGAGTTGCACTCAAC 1290
Db 1081 TCCACCGCTTCGGTGAGACCGAGGACCAACCAATTCGCTCAGCTTCGAGTTGCACTCAAC 1140

Qy 1291 TGTGGCAGATCAAGACTGGTGTCCAGACGTTTCGACCGGTGTTCGAAAGTCAACACGAG 1350
Db 1141 TGTGGCAGATCAAGACTGGTGTCCAGACGTTTCGACCGGTGTTCGAAAGTCAACACGAG 1200

Qy 1351 CTTCTCCGCAATCGAGAGCTGCTTTGGCGAGCGCGGCGTCTACGAGGCTCGAGCGCATTC 1410
Db 1201 CTTCTCCGCAATCGAGAGCTGCTTTGGCGAGCGCGGCGTCTACGAGGCTCGAGCGCATTC 1260

```

```

Qy 1411 CCACGCTTTTCAGGCTAA 1428
Db 1261 CCACGCTTTTCAGGCTAA 1278

RESULT 6
ID AAH66050
XX AAH66050;
AC AC
XX DT 26-SEP-2001 (first entry)
XX DE C glutamicum coding sequence fragment SEQ ID NO: 1085.
XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX OS organic acid synthesis; ds.
XX PN Corynebacterium glutamicum.
XX PF BPI108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX DR WPI; 2001-376931/40.
XX DR P-PSDB; AAG90831.

Novel polynucleotides derived from Corynebacterium bacteria, for identifying
mutation point of a gene, measuring expression of a gene, analyzing
expression profile or pattern of a gene and identifying homologous gene.

Claim 8; SEQ ID NO 1085; 246pp + Sequence Listing; English.

The present invention provides a number of nucleotide and protein
sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
are useful for identifying the mutation point of a gene derived from a
mutant of corynebacterium bacterium, measuring expression amount and analyzing
the expression profile or expression pattern of a gene derived from
Corynebacterium bacterium, and identifying a homologue of a gene derived from
Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
acids, nucleic acids, vitamins, saccharides and organic acids,
particularly L-lysine. The present sequence is a nucleic acid described
in the exemplification of the invention. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from the European Patent Office

Sequence 1275 BP; 269 A; 399 C; 347 G; 260 T; 0 U; 0 Other;

Query Match 80.7%; Score 1273.4; DB 5; Length 1275;
Best Local Similarity 99.9%; Pred. No. 1.7e-312;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 151 GTGGCTGAAATCATGACGATTCGCTCGGAAATTCGACTCCCGCGGTAAACCAACC 210
Db 1 GTGGCTGAAATCATGACGATTCGCTCGGAAATTCGACTCCCGCGGTAAACCAACC 60

Qy 211 GTGAGGCGAGAGTTTTCCTGGATGACGTTCCCGAGTTCGAGGTTCATCCGCGC 270
Db 61 GTGAGGCGAGAGTTTTCCTGGATGACGTTCCCGAGTTCGAGGTTCATCCGCGC 120

Qy 271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCTGACGCTGCGAGTTCGCTCGGC 330
Db GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCTGACGCTGCGAGTTCGCTCGGC

```

121 GCATCCACCGGGCTCCAGAGGCTCATGAGTGGTGACGGTGGCGATCGCTACCTGGGC 180
331 AAGGCGCTTTGAAGGCGAGTTGAAAACGTCAA CGAAGAAATCGCGCAGCGCTCGTGGC 390
181 AAGGCGCTTTGAAGGCGAGTTGAAAACGTCAA CGAAGAAATCGCGCAGCGCTCGTGGC 240
391 CTAGAGGCTCAGCATCAGCGCTCATCGAGAGCAATGATCAGCTTGATGGCAGCGCC 450
241 CTAGAGGCTCAGCATCAGCGCTCATCGAGAGCAATGATCAGCTTGATGGCAGCGCC 300
451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGTTGTTTCCATGGCTGTGCAAAAGGCT 510
301 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGTTGTTTCCATGGCTGTGCAAAAGGCT 360
511 GCTGCTGATTTCCGAGGCTCCCATCTGTTCCGTACATCCGGTGGACCAACGCAACGCTT 570
361 GCTGCTGATTTCCGAGGCTCCCATCTGTTCCGTACATCCGGTGGACCAACGCAACGCTT 420
571 CTTCCAGTTTCCAATGATGAACATCATCAACGGTGGGCTCAGCGTGAATCCGGTGTGAC 630
421 CTTCCAGTTTCCAATGATGAACATCATCAACGGTGGGCTCAGCGTGAATCCGGTGTGAC 480
631 GTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC 690
481 GTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC 540
691 GGGCGGAGGTTTACACGCACTGAAGTCCGTGATCAAGGAAAGGGCCCTGTCCACGGGA 750
541 GGGCGGAGGTTTACACGCACTGAAGTCCGTGATCAAGGAAAGGGCCCTGTCCACGGGA 600
751 CTTGGCGATCAGGCGGCTTTCGCTTTCGTTCCGTGGCTCCACCGTGAAGGCTCTTGACCTT 810
601 CTTGGCGATCAGGCGGCTTTCGCTTTCGTTCCGTGGCTCCACCGTGAAGGCTCTTGACCTT 660
811 ATCGTTAAGCAATTCGAGAAGGCTGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTG 870
661 ATCGTTAAGCAATTCGAGAAGGCTGCTTCAACCCAGGCAAGGACATCGCTCTTGCTCTG 720
871 GAGTTGCTTTCCTGAGTTTCTCAAGGAGCGCACTTACACTTCGAGGTGGCGAGCAC 930
721 GAGTTGCTTTCCTGAGTTTCTCAAGGAGCGCACTTACACTTCGAGGTGGCGAGCAC 780
931 TCCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGGGTACCCCAATCGTCTCC 990
781 TCCGAGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGGGTACCCCAATCGTCTCC 840
991 ATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTTACACCAACCTTACCGCAACCATC 1050
841 ATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTTACACCAACCTTACCGCAACCATC 900
1051 GGGCAGAGGTTTCAAGTCTGTTGGGAGGACTTCTTGTACCAACCTTGAAGCGCTGAAG 1110
901 GGGCAGAGGTTTCAAGTCTGTTGGGAGGACTTCTTGTACCAACCTTGAAGCGCTGAAG 960
1111 GAGGGATCGCTAAGAGGCTGCCAATCCATCTGTTTAAAGGTGAACCAAGATCGGTACC 1170
961 GAGGGATCGCTAAGAGGCTGCCAATCCATCTGTTTAAAGGTGAACCAAGATCGGTACC 1020
1171 CTCACGAGACCTTCGACGCTGCGACATCGGCTCACCGCGAGGCTACACCTCCATGATG 1230
1021 CTCACGAGACCTTCGACGCTGCGACATCGGCTCACCGCGAGGCTACACCTCCATGATG 1080
1231 TCCACCGTTTCGGTGAAGCGGAGGACACCACTTGTGACCTTCGAGTTGCACTCAAC 1290
1081 TCCACCGTTTCGGTGAAGCGGAGGACACCACTTGTGACCTTCGAGTTGCACTCAAC 1140
1291 TGTGGCCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGTGTCCGAAAGTACACCGAG 1350
1141 TGTGGCCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGTGTCCGAAAGTACACCGAG 1200
1351 CTTCTCCGATTCGAGCAGCTGCTTTGGCGACCGCGGCGTCTACGAGGTTCGACGGCATTC 1410
1201 CTTCTCCGATTCGAGCAGCTGCTTTGGCGACCGCGGCGTCTACGAGGTTCGACGGCATTC 1260

Qy 1411 CCACGCTTTCAGGGC 1425
Db 1261 CCACGCTTTCAGGGC 1275

RESULT 7

ACC45312
ID ACC45312 standard; DNA; 1299 BP.

XX ACC45312;

XX 17-JUN-2003 (first entry)

XX Corynebacterium glutamicum eno DNA fragment SEQ ID NO:3.

XX Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
KW genetically modified microorganism; metabolite; biosynthesis; amino acid;
KW vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
KW pharmaceutical; food; animal feeding; eno; gene; ds.

XX Corynebacterium glutamicum.

OS

XX WO2003023016-A2.

PN

XX 20-MAR-2003.

XX 11-SEP-2002; 2002WO-EP010174.

PF

XX 13-SEP-2001; 2001DE-01045043.

PR

XX (DEGS) DEGUSSA AG.

XX Farwick M, Hermann T;

XX WPI; 2003-354534/33.

XX Microorganism useful for producing e.g. fine chemicals, has permanently

PT altered phosphorylatability protein, such that biosynthesis of fine

PT chemical synthesized by microorganism is increased compared to wild-type.

XX Example 3; Page 81-82; 120pp; English.

XX The present invention describes a microorganism (I), in which the

XX phosphorylatability of at least one protein has been permanently altered

XX such that the biosynthesis of at least one fine chemical synthesised by

XX the microorganism is increased compared to the wild type. Also described:

XX (1) use of a DNA (II) sequence coding for a protein which contains a

XX phosphorylation site, where the sequence contains such a mutation that

XX the protein is changed in its phosphorylatability for the production of

XX (I), or for the production of fine chemicals; and (2) a method for

XX producing fine chemicals or metabolites comprising using (I). (I) is

XX useful for producing fine chemicals or metabolites, such as amino acids,

XX vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids

XX and vitamins produced using (I) can be used in human medicine. (I)

XX pharmaceutical industry, food industry and in animal feeding. (I)

XX produces larger amount of desired fine chemical or a metabolite than the

XX wild type. The present sequence represents a wild type eno fragment DNA

XX sequence from Corynebacterium glutamicum, which is used in an example

XX from the present invention

XX SQ Sequence 1299 BP; 274 A; 409 C; 353 G; 263 T; 0 U; 0 Other;

Query Match 80.7%; Score 1273.4; DB 8; Length 1299;

Best Local Similarity 99.9%; Pred. No. 1.7e-312;

Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 154 GCTGAATCATGACGCTATTTCGCTCCGAAATTCGACTCCCGCGGTACCCACCGTC 213

Db 10 GCTGAATCATGACGCTATTTCGCTCCGAAATTCGACTCCCGCGGTACCCACCGTC 69

Qy 214 GAGGCGAGGTTTTCTTGGATGACGGTCCCGCGGTGTCCACGGTGTCCATCCGCGCA 273

Db 70 GAGCAGAGGTTTCTCGGATGACGGTTCCACGGTGTGCGAGGTGTTCATCCGGCGCA 129
Qy 274 TCCACCGGCTCCACGAGGCTCATGAGCTGGTGACGGTGGCATCGCTACTCGGCAAG 333
Db 130 TCCACCGGCTCCACGAGGCTCATGAGCTGGTGACGGTGGCATCGCTACTCGGCAAG 189
Qy 334 GGCGTTTGAAGGCGATTGAAACGTTCAACGAAAGATCGCGACGAGCTGCTGGGCTA 393
Db 190 GGCGTTTGAAGGCGATTGAAACGTTCAACGAAAGATCGCGACGAGCTGCTGGGCTA 249
Qy 394 GAGGCTGACGATGAGCGCTCATCGAGNAGCAATGATCAAGCTTGATGCGACCGCAAC 453
Db 250 GAGGCTGACGATGAGCGCTCATCGAGNAGCAATGATCAAGCTTGATGCGACCGCAAC 309
Qy 454 AAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGCAAGGCTGCT 513
Db 310 AAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTGCAAGGCTGCT 369
Qy 514 GCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAACCAACGACACGTTCTT 573
Db 370 GCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGAACCAACGACACGTTCTT 429
Qy 574 CCAGTCCCAATGATGAACATCAACGCTGGCGCTCACGCTGACTCGGTGTTGACGTT 633
Db 430 CCAGTCCCAATGATGAACATCAACGCTGGCGCTCACGCTGACTCGGTGTTGACGTT 489
Qy 634 CAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
Db 490 CAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGC 549
Qy 694 GCGGAGTCTACCAACGACATGAAGTCCGCTCATCAAGGAAAGGCGCTGTCCACCGGACTT 753
Db 550 GCGGAGTCTACCAACGACATGAAGTCCGCTCATCAAGGAAAGGCGCTGTCCACCGGACTT 609
Qy 754 GCGGATGAGGCGGCTTCGCTCTCGTCCGTCGCTGCGCTCCACCGTGAAGCTCTTGACCTTATC 813
Db 610 GCGGATGAGGCGGCTTCGCTCTCGTCCGTCGCTCCACCGTGAAGCTCTTGACCTTATC 669
Qy 814 GTTAAGGCAATCGAAGAGGCTGGCTTCAACCGGCAAGGACATCGCTCTTGCTCTCGAC 873
Db 670 GTTAAGGCAATCGAAGAGGCTGGCTTCAACCGGCAAGGACATCGCTCTTGCTCTCGAC 729
Qy 874 GTTGCTTCTCTGAGTTCTTCAAGGACGGACCTTACCACCTTCGAAGGTGGCCAGCACTCC 933
Db 730 GTTGCTTCTCTGAGTTCTTCAAGGACGGACCTTACCACCTTCGAAGGTGGCCAGCACTCC 789
Qy 934 GCAGCTGAGATGGCAAAAGTTTACGCTGAGCTGCTGAGCGGTACCAATCGTCTCCATC 993
Db 790 GCAGCTGAGATGGCAAAAGTTTACGCTGAGCTGCTGAGCGGTACCAATCGTCTCCATC 849
Qy 994 GAGGACCCACTGACAGGAAGATGACTGGGAGGTTTACACCAACCTCACCGCAACCATCGC 1053
Db 850 GAGGACCCACTGACAGGAAGATGACTGGGAGGTTTACACCAACCTCACCGCAACCATCGC 909
Qy 1054 GACAAGGTTGAGATCGTTGGCGAGCTTCTTCGTCACCAACCTTGAGCGCTGAAGGAG 1113
Db 910 GACAAGGTTGAGATCGTTGGCGAGCTTCTTCGTCACCAACCTTGAGCGCTGAAGGAG 969
Qy 1114 GGATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACAGATCGGTACCCCTC 1173
Db 970 GGATCGCTAAGAGGCTGCCAATCCATCTGTTAAGGTGAACAGATCGGTACCCCTC 1029
Qy 1174 ACCGAGACCTTCGACGCTGTGCATGCTCACCGCGAGGTACCTCCATGATGCTCC 1233
Db 1030 ACCGAGACCTTCGACGCTGTGCATGCTCACCGCGAGGTACCTCCATGATGCTCC 1089
Qy 1234 CACGTTCCGCTGAGACCGAGGACACCACTTCTGACCTCGAGTTGCACTCAACTGT 1293
Db 1090 CACGTTCCGCTGAGACCGAGGACACCACTTCTGACCTCGAGTTGCACTCAACTGT 1149
Qy 1294 GGCAGATCAAGACTGCTGCTCCAGCAAGTTCGACGCTGTCGAAAGTACAAACGACTT 1353
Db 1150 GGCAGATCAAGACTGCTGCTCCAGCAAGTTCGACGCTGTCGAAAGTACAAACGACTT 1209

Qy 1354 CTCGCATCGAGCAGCTGCTTGGCGACCGCGGCTCTACGAGGTCCGAGCGCATTTCCCA 1413
Db 1210 CTCGCATCGAGCAGCTGCTTGGCGACCGCGGCTCTACGAGGTCCGAGCGCATTTCCCA 1269
Qy 1414 CGCTTTCAGGCTTAA 1428
Db 1270 CGCTTTCAGGCTTAA 1284
RESULT 8
ACC45314
ID ACC45314 standard; DNA; 1278 BP.
XX AC ACC45314;
XX AC ACC45314;
DT 17-JUN-2003 (first entry)
XX Modified eno S330E encoding DNA SEQ ID NO:5.
DE Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;
XX genetically modified microorganism; metabolite; biosynthesis; amino acid;
XX vitamin; nucleoside; nucleotide; pigment; protein; human medicine;
XX pharmaceutical; food; animal feeding; eno; gene; ds.
OS Corynebacterium glutamicum.
OS Synthetic.
XX Key Location/Qualifiers
XX CDS 1..1278
XX FT /*tag= a
XX FT /product= "modified eno S330E protein"
XX PN WO2003023016-A2.
XX 20-MAR-2003.
XX PD 11-SEP-2002; 2002WO-EP010174.
XX PF 13-SEP-2001; 2001DE-01045043.
XX PR (DEGS) DEGUSSA AG.
XX Farwick M, Hermann T;
XX WPI; 2003-354534/33.
XX P-PSDB; ABP97002.
XX Microorganism useful for producing e.g. fine chemicals, has permanently
XX altered phosphorylatability protein, such that biosynthesis of fine
XX chemical synthesized by microorganism is increased compared to wild-type.
XX Example 3; Page 83-84; 120pp; English.
XX The present invention describes a microorganism (I), in which the
XX phosphorylatability of at least one protein has been permanently altered
XX such that the biosynthesis of at least one fine chemical synthesised by
XX the microorganism is increased compared to the wild type. Also described:
XX (1) use of a DNA (II) sequence coding for a protein which contains a
XX phosphorylation site, where the sequence contains such a mutation that
XX the protein is changed in its phosphorylatability for the production of
XX (1), or for the production of fine chemicals; and (2) a method for
XX producing fine chemicals or metabolites comprising using (1). (1) is
XX useful for producing fine chemicals or metabolites, such as amino acids,
XX vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids
XX and vitamins produced using (1) can be used in human medicine, in the
XX pharmaceutical industry, food industry and in animal feeding. (1)
XX produces larger amount of desired fine chemical or a metabolite than the
XX wild type. The present sequence encodes a modified eno S330E from
XX Corynebacterium glutamicum, which is used in an example from the present
XX invention
XX Sequence 1278 BP; 272 A; 397 C; 349 G; 260 T; 0 U; 0 Other;

Query Match				80.6%; Score 1271.6; DB 8; Length 1278;			
Best Local Similarity				99.7%; Pred. No. 4.8e-312;			
Matches 1274; Conservative				0; Mismatches 312; Indels 0; Gaps 0;			
Qy	151	GTGCTGAAATCATATGACGCTATTCGCTCGGGAATTCCTCGACTCCCGCGGTAAACCCAAACC	210				
Db	1	GTGCTGAAATCATATGACGCTATTCGCTCGGGAATTCCTCGACTCCCGCGGTAAACCCAAACC	60				
Qy	211	GTGAGGACAGGTTTCTTGATGACGGTTCCACCGGTGTGCGAGGTGTTCATCCGCG	270				
Db	61	GTGAGGACAGGTTTCTTGATGACGGTTCCACCGGTGTGCGAGGTGTTCATCCGCG	120				
Qy	271	GCATCCACCGCGTCCACGAGGCTCATGAGTGGTGAACGCTCAACGAGAAATCGCGACGAGCTCGCTGGC	330				
Db	121	GCATCCACCGCGTCCACGAGGCTCATGAGTGGTGAACGCTCAACGAGAAATCGCGACGAGCTCGCTGGC	180				
Qy	331	AAGGGCGTTTGAAGGCAGTTGAAACGCTCAACGAGAAATCGCGACGAGCTCGCTGGC	390				
Db	181	AAGGGCGTTTGAAGGCAGTTGAAACGCTCAACGAGAAATCGCGACGAGCTCGCTGGC	240				
Qy	391	CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGCACCGCC	450				
Db	241	CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGATGGCACCGCC	300				
Qy	451	AACAAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGGCTGTTGCAAAAGGCT	510				
Db	301	AACAAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTTCATGGCTGTTGCAAAAGGCT	360				
Qy	511	GCTGCTGATTCGCGAGGCTCCCACTGTTCGCTACATCGGTGGACCAACGACACAGTT	570				
Db	361	GCTGCTGATTCGCGAGGCTCCCACTGTTCGCTACATCGGTGGACCAACGACACAGTT	420				
Qy	571	CTTCCAGTTCGAATGATGAACATCATCAAGGTGGCGCTCAGCTGATCCGGTGTGAC	630				
Db	421	CTTCCAGTTCGAATGATGAACATCATCAAGGTGGCGCTCAGCTGATCCGGTGTGAC	480				
Qy	631	GTTCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	690				
Db	481	GTTCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC	540				
Qy	691	GGCGCGAGTCTACCAACGACATGAAGTCCGTATCAAGGAAAGGGCTGTCCACCGGA	750				
Db	541	GGCGCGAGTCTACCAACGACATGAAGTCCGTATCAAGGAAAGGGCTGTCCACCGGA	600				
Qy	751	CTTGGCGATGAGGCGGCTTCGCTCTTCGCTGGCTTCCACCGTGAGGCTCTTGACCTT	810				
Db	601	CTTGGCGATGAGGCGGCTTCGCTCTTCGCTGGCTTCCACCGTGAGGCTCTTGACCTT	660				
Qy	811	ATGTTTAAAGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTTGTGCTCG	870				
Db	661	ATGTTTAAAGCAATCGAAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTTGTGCTCG	720				
Qy	871	GAGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTTACCACTTCGAAGGTGGCCAGCAC	930				
Db	721	GAGTTGCTTCTCTGAGTTCTTCAAGGACGGCACCTTACCACTTCGAAGGTGGCCAGCAC	780				
Qy	931	TCCGACGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGGGTACCAATCGTCTCC	990				
Db	781	TCCGACGCTGAGATGGCAACGTTTACGCTGAGCTGTTGACGGGTACCAATCGTCTCC	840				
Qy	991	ATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTTACACCAACCTTCCCGCAACCATC	1050				
Db	841	ATCGAGGACCCACTGCGAGGAAGATGACTGGGAGGGTTTACACCAACCTTCCCGCAACCATC	900				
Qy	1051	GGCGACAAGTTGATGCTGTTGGGACGATTTCTTGTGTCACCAACCTTGAGCGGCTGAAG	1110				
Db	901	GGCGACAAGTTGATGCTGTTGGGACGATTTCTTGTGTCACCAACCTTGAGCGGCTGAAG	960				
Qy	1111	GAGGGATCGCTAAGAGGCTGCCAATCCATCCTGTTAAGGTGAACAGATCGGTACC	1170				
Db	961	GAGGGATCGCTAAGAGGCTGCCAACGAGATCCTGTTAAGGTGAACAGATCGGTACC	1020				

Qy	1171	CTCACGAGACCTTCGACGCTGTCGACATGGCTCACGGCAGGCTACACCTCCATGATG	1230
Db	1021	CTCACGAGACCTTCGACGCTGTCGACATGGCTCACGGCAGGCTACACCTCCATGATG	1080
Qy	1231	TCCACCGCTTCCGGTGAGACCGGAGGACACCCACCATTTGCTGACCTCGCAGTTGCACTCAAC	1290
Db	1081	TCCACCGCTTCCGGTGAGACCGGAGGACACCCACCATTTGCTGACCTCGCAGTTGCACTCAAC	1140
Qy	1291	TGTGGCCAGATCAAGACTGTGTCTCAGACACGTTCCGACCCGTGTGCAAAAGTACAAACAG	1350
Db	1141	TGTGGCCAGATCAAGACTGTGTCTCAGACACGTTCCGACCCGTGTGCAAAAGTACAAACAG	1200
Qy	1351	CTTCTCCGATCCAGCAGCTGCTTGGCGAGCGCGGCTCTACGACGTCGACGGCATTC	1410
Db	1201	CTTCTCCGATCCAGCAGCTGCTTGGCGAGCGCGGCTCTACGACGTCGACGGCATTC	1260
Qy	1411	CCACGCTTTTCAGGGCTAA 1428	
Db	1261	CCACGCTTTTCAGGGCTAA 1278	
RESULT 9			
ACC45313			
ID	ACC45313 standard; DNA; 1299 BP.		
XX			
AC	ACC45313;		
XX			
DT	17-JUN-2003 (first entry)		
XX	Modified eno DNA fragment S330E SEQ ID NO:4.		
DE			
XX	Fine chemical; Coryneform bacteria; Escherichia coli; microorganism;		
KW	genetically modified microorganism; metabolite; biosynthesis; amino acid;		
KW	vitamin; nucleoside; nucleotide; pigment; protein; human medicine;		
KW	pharmaceutical; food; animal feeding; eno; gene; ds.		
XX			
OS	Corynebacterium glutamicum.		
OS	Synthetic.		
XX			
PN	W02003023016-A2.		
XX			
PD	20-MAR-2003.		
XX			
PF	11-SEP-2002; 2002WO-EP010174.		
XX			
PR	13-SEP-2001; 2001DE-01045043.		
XX			
PA	(DEGS) DEGUSSA AG.		
XX			
PI	Farwick M, Hermann T;		
XX			
DR	WPI; 2003-354534/33.		
XX			
PT	Microorganism useful for producing e.g. fine chemicals, has permanently		
PT	altered phosphorylatability protein, such that biosynthesis of fine		
PT	chemical synthesized by microorganism is increased compared to wild-type.		
XX			
PS	Claim 17; Page 82; 120pp; English.		
XX			
CC	The present invention describes a microorganism (I), in which the		
CC	phosphorylatability of at least one protein has been permanently altered		
CC	such that the biosynthesis of at least one fine chemical synthesised by		
CC	the microorganism is increased compared to the wild type. Also described:		
CC	(1) use of a DNA (II) sequence coding for a protein which contains a		
CC	phosphorylation site, where the sequence contains such a mutation that		
CC	the protein is changed in its phosphorylatability for the production of		
CC	(I), or for the production of fine chemicals; and (2) a method for		
CC	producing fine chemicals or metabolites comprising using (I). (I) is		
CC	useful for producing fine chemicals or metabolites, such as amino acids,		
CC	vitamins, nucleosides, nucleotides, pigments or proteins. The amino acids		
CC	and vitamins produced using (I) can be used in human medicine, in the		
CC	pharmaceutical industry, food industry and in animal feeding. (I)		
CC	produces larger amount of desired fine chemical or a metabolite than the		

CC wild type. The present sequence represents a modified eno fragment DNA
CC sequence from Corynebacterium glutamicum, which is used in an example
CC from the present invention
XX
SQ Sequence 1299 BP; 275 A; 407 C; 355 G; 262 T; 0 U; 0 Other;

Query Match 80.4%; Score 1268.6; DB 8; Length 1299;
Best Local Similarity 99.7%; Pred. No. 2.8e-311;
Matches 1271; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 154 GCTGAATCATGACAGCTATTCCGTCGCGAAATTCCTCGACTCCCGCGGTAAACCCACCGTC 213
Db 10 GCTGAATCATGACAGCTATTCCGTCGCGAAATTCCTCGACTCCCGCGGTAAACCCACCGTC 69

Qy 214 GAGCAGAGGTTTCCCTGGATGACCGTTTCCCAAGGTTTCCATCCGCGGCA 273
Db 70 GAGCAGAGGTTTCCCTGGATGACCGTTTCCCAAGGTTTCCATCCGCGGCA 129

Qy 274 TCCACCGCGTCCACGAGGCTCATGAGTGGTGACCGTGGCGATCGTACCTGGGCAAG 333
Db 130 TCCACCGCGTCCACGAGGCTCATGAGTGGTGACCGTGGCGATCGTACCTGGGCAAG 189

Qy 334 GCGGTTTGAAGGCGAGTTGAAGAGCTCAACGAGAAATTCGCGACGAGCTCGCTGGCCTA 393
Db 190 GCGGTTTGAAGGCGAGTTGAAGAGCTCAACGAGAAATTCGCGACGAGCTCGCTGGCCTA 249

Qy 394 GAGGCTGACGATCAGCGCTCATCGACGAAAGCAATGATCAAGCTTGATGGCACCGCAAC 453
Db 250 GAGGCTGACGATCAGCGCTCATCGACGAAAGCAATGATCAAGCTTGATGGCACCGCAAC 309

Qy 454 AAGTCCCGCTGGTGCAAGCGAAATCTTGGTGTTCATGGCTGTGCAAAAGGCTGCT 513
Db 310 AAGTCCCGCTGGTGCAAGCGAAATCTTGGTGTTCATGGCTGTGCAAAAGGCTGCT 369

Qy 514 GCTGATTCGCGAGGCTCCCACTGTCGCTGACATCGGTGACCGTGGACCAACGACAGTTCCT 573
Db 370 GCTGATTCGCGAGGCTCCCACTGTCGCTGACATCGGTGACCGTGGACCAACGACAGTTCCT 429

Qy 574 CCAGTTCCAATGATGAACATCATCAAGCGTGGCGCTCACGCTGACTCCGGTGTGACGTT 633
Db 430 CCAGTTCCAATGATGAACATCATCAAGCGTGGCGCTCACGCTGACTCCGGTGTGACGTT 489

Qy 634 CAGGAATTCATGATCGCTTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGC 693
Db 490 CAGGAATTCATGATCGCTTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAACGGC 549

Qy 694 GCGAGTCTACCAACGACCTGAAGTCCGTATCAAGGAAAGGCGCTGTCAACGGAATT 753
Db 550 GCGAGTCTACCAACGACCTGAAGTCCGTATCAAGGAAAGGCGCTGTCAACGGAATT 609

Qy 754 GCGATGAGGCGGCTTCCGCTTCCGTCGCTCCACCGTCCACCGTGGAGGCTCTTGACCTTATC 813
Db 610 GCGATGAGGCGGCTTCCGCTTCCGTCGCTCCACCGTGGAGGCTCTTGACCTTATC 669

Qy 814 GTTAAGCAATCGAAGAGGCTGGCTTCAACCCGAGGCAAGGACATCGCTTGTCTGGAAC 873
Db 670 GTTAAGCAATCGAAGAGGCTGGCTTCAACCCGAGGCAAGGACATCGCTTGTCTGGAAC 729

Qy 874 GTTGCTTCTCTGAGTTCCTCAAGGACGCGACCTTACCACTTCGAAGTGGCGACATCC 933
Db 730 GTTGCTTCTCTGAGTTCCTCAAGGACGCGACCTTACCACTTCGAAGTGGCGACATCC 789

Qy 934 GCAGCTGAGATGCAACAGTTTACGCTGAGCTGTTGACCGGTACCAATCGTCTCCATC 993
Db 790 GCAGCTGAGATGCAACAGTTTACGCTGAGCTGTTGACCGGTACCAATCGTCTCCATC 849

Qy 994 GAGGACCCACTGCGAGGAAGTACTGCGAGGTTTACACCAACCTCACCGCAACCATCGG 1053
Db 850 GAGGACCCACTGCGAGGAAGTACTGCGAGGTTTACACCAACCTCACCGCAACCATCGG 909

Qy 1054 GACAGGTTGAGTGGTGGCGAGCTTCTTGGTCAACCAACCTGAGCGCCTGAAGGAG 1113
Db 910 GACAGGTTGAGTGGTGGCGAGCTTCTTGGTCAACCAACCTGAGCGCCTGAAGGAG 969

Qy 1114 GGATCGCTTAAGAGGCTGCAACTCCATCCTGTTAAGGTGAACACGATCGGTACCCCTC 1173
Db 970 GGATCGCTTAAGAGGCTGCAACTCCATCCTGTTAAGGTGAACACGATCGGTACCCCTC 1029

Qy 1174 ACCGAGACCTTCGACGCTGCGATGCTGACATGCTCACCGCGAGGCTACACTTCATGATGCC 1233
Db 1030 ACCGAGACCTTCGACGCTGCGATGCTGACATGCTCACCGCGAGGCTACACTTCATGATGCC 1089

Qy 1234 CACGTTCCGCTGAGACCGAGGACACCACTTCTGACCTTCGAGTTGCACTCAACTGT 1293
Db 1090 CACGTTCCGCTGAGACCGAGGACACCACTTCTGACCTTCGAGTTGCACTCAACTGT 1149

Qy 1294 GGCCAGATCAAGACTGGTGTCTCCAGCACTGTTCCGACCGTGTGCAAAAGTCAACACGCTT 1353
Db 1150 GGCCAGATCAAGACTGGTGTCTCCAGCACTGTTCCGACCGTGTGCAAAAGTCAACACGCTT 1209

Qy 1354 CTCCGATCGACGAGCTGCTTGGCGACGCGCGCTCTACGAGGTGCGAGGCAATTCCTCA 1413
Db 1210 CTCCGATCGACGAGCTGCTTGGCGACGCGCGCTCTACGAGGTGCGAGGCAATTCCTCA 1269

Qy 1414 CGCTTTCAGGCTAA 1428
Db 1270 CGCTTTCAGGCTAA 1284

RESULT 10
ACA29854
ID ACA29854 standard; DNA; 1275 BP.
XX ACA29854;
AC AC
XX AC
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #11511.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Corynebacterium diphtheriae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 08-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR P-PSDB; ABU25984.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 17724; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 5213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
CC antiseptic nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antiseptic nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
XX
SQ Sequence 1275 BP; 268 A; 378 C; 354 G; 275 T; 0 U; 0 Other;

Query Match 57.5%; Score 907; DB 8; Length 1275;
Best Local Similarity 82.0%; Pred. No. 1.4e-219;
Matches 1045; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

QY 151 GTGGCTGAATCATGACGATTCGCTCGCGAAATTCCTGCACTCCCGCGGTAAACCCAAAC 210
DB 1 GTGGCTGACATTAATGACGATTTTGTCTGTGAATCTTGACCTCCCGCGGTAAACCCAAAC 60

QY 211 GTGAGGCGACAGGTTTCTCGGATGACGGTTCCACGGTGTCCGAGGTGTTCATCCGGC 270
DB 61 GTTGAAGCCGAAGTTTCTTGGATGACGATCCACGGCGTTGGGGCGTTCTTCCGGT 120

QY 271 GCATCCACCGCGTCCACGAGGCTCATGAGTGTGCGTACGCGTGGCGATCGCTACCTGGGC 330
DB 121 GCGTCCACCGGTTTCACGAGGCTCACGAGCTTCGCGACGGTGGCGAGCGCTACCTCGGC 180

QY 331 AAGGGCGTTTGAAGGCGATTGAAACGTCAAGCAAAATTCGCGACGAGCTCGCTGGC 390
DB 181 AAGGGCGTTTCAATGAGTAAACACGTGAACGAGGAAATCGCTGAACCAATCGCTGGT 240

QY 391 CTAGAGGCTGACGATCAGCGCTCATCGAAGCAATGATCAAGCTTGTATGGCACCGCC 450
DB 241 GCAGAACCCAGCATCAGCGCTGATCGATCAGCGGATGATGCTCTCGACGGCAGCTGAG 300

QY 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGGCTGTTCGAAAGGCT 510
DB 301 AACAAAGTCTCGTCTCGCGCAACGCTATCTTGGGTGTATCCATCCGCGGTAGCTAAGGCT 360

QY 511 GCTGCTGATTCGCGAGCGCTCCCACTGTTCGCTACATCGGTGGAACCAACGACAGCTT 570
DB 361 GCGCGAGAGTCTGCTGGCTACCTTTGTACGCTTACATCGCGCGCCCTAACGCTTCAGTT 420

QY 571 CTTCCAGTCCCAATGATGAACATCATCAAGGTGGCGCTCACGCTGACTCGGTGTTGAC 630
DB 421 CTTCCAGTCTTATGATGAACATTTGTAAGGTGGCGCACGCTGACTCGGTGTTGAT 480

QY 631 GTTCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAAC 690
DB 481 GTTCAGGAGTTTATGATGCTCTTATCGGTGCCGAGTCTTCTCTGAGGCTCTGGCATG 540

QY 691 GGCGCGAGGTTCTACCAAGCACTGAAGTCCGTCATCAAGGAAAGGCGCTGTCCACCGGA 750
DB 541 GGTGCAGAGGTTCTACCACTCTTGAAGTCCGTCATTAAGTCCAAAGGAGCTTTCCACCGGC 600

QY 751 CTTGGCGATGAGGCGGCTTCTGCTTCTTCCGTCGGCTCCACCGCTGAGGCTCTTGACCTT 810

DB 601 CTCGGCGACGAAGGTGTTTCGCACCTTCTGTGTAGTCCACCAAGCAGCTCTCGACCTC 660
QY 811 ATCGTTAAAGCAATCGAAGAGGTGCTGTCAACCCAGGCAAGGACATCGCTTGTGCTGTG 870
DB 661 ATCGTTGAGGCAATTCGAAGAGGTGCTGTCAAGCCAGGTGCTGACATTCGCACTTGCAC 720
QY 871 GAGGTTGCTTCTCTGAGTTCCTCAAGGACGCGACCTTACACCTTCGCAAGGTGSCCAGCAC 930
DB 721 GAGGTTGCTTCTCTGAGTTCCTCAAGGATGGAAGTACACCTTCGAGGGGGGCGAGCAC 780
QY 931 TCCGACAGCTGAGATGTCACAAACGTTTACGCTGAGCTGTTGACGCGTACCCAAATCGTCTCC 990
DB 781 ACGCTGAGGAGATGCAAGGCTTACGAGCAGCTCATCGTGAGTACCAATTTGTTTCC 840
QY 991 ATCGAGGACCCACTGTCAGGAAGATGCTGGAGGGTTTACCAACCTCACCGCAACATC 1050
DB 841 ATCGAGGACCCACTGTCAGGAAGACGACTGGAGGGCTTACACCGCTGACCGCGCAATC 900
QY 1051 GCGCAAGAGGTTTCAGATCGTTGCGGACGACTTCTTCTGTCACCAACCTGAGCGCTGAAG 1110
DB 901 GGTGCAAGGTTTCAGATCGTTGCGGACGACTTCTTCTGTCACCAACCTGAGCGCTCAAG 960
QY 1111 GAGGCGATCGCTAAGAAAGGCTGCCAACTCCATCTCTGTTAAAGGTGAACAGATCGGTACC 1170
DB 961 GAAGGATCGAGAAGAGGCTGCCAATGCTTGTCTGTTAAAGGTCAACAGATCGGTACC 1020
QY 1171 CTCACCGAGACCTTCGACGCTGTGCATGGCTCACCGCGAGGCTTACACCTCCATGATG 1230
DB 1021 CTGACCGAGACCTTCGACGCTGTGATCTCGCACACCGCAACCGCTTACCGCACCATG 1080
QY 1231 TCCACCGGTTCCGGTTCAGACCGGAGACACCACTTGTGACCTCCGAGTTGCACTCAAC 1290
DB 1081 TCCACCGGTTCCGGTTCAGACCGGAGACTGAAGACACCACTTGTGACCTTGCAGTCGAT 1140
QY 1291 TGTGGCAGATCAAGACTGCTGCTCCAGCACGTTCCGACCGTGTCCGAAAGTACAACAC 1350
DB 1141 TGTGGCAGATCAAGACCGGTCGACAGCTGTTCCGAGCGGCTTGCAGAGTCAATCAG 1200
QY 1351 CTTCTCCGATCGAGCAGCTGCTTTCGCGACGCGCGCTCTACGCGGTTCGACGCGATTC 1410
DB 1201 CTTCTCCGATCGAGCAGCAGCTTTCGATGTCAGCAGTCTTACGCGAGCGCTTTCAGCAT 1260
QY 1411 CCACGCTTTCAGGCG 1425
DB 1261 CCACGCTTTCAGGCG 1275

RESULT 11
AD56353
ID AD56353 standard; cDNA; 1269 BP.
XX
AC AD56353;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #8340.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.
XX Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.
 PR (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PT
 XX
 PS Claim 1; SEQ ID NO 32027; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1269 BP; 227 A; 421 C; 443 G; 178 T; 0 U; 0 Other;
 Query Match 43.0%; Score 678.2; DB 13; Length 1269;
 Best Local Similarity 71.0%; Pred. No. 1.4e-161;
 Matches 899; Conservative 0; Mismatches 368; Indels 0; Gaps 0;
 QY 152 TGGCTGAAATCATGCAGGATTCGCTCGGAAATTCGACTCCCGGGTAAACCCACCG 211
 DB 2 TGGCTGATCGAGGCGAGTACGCGACGGAGATTCGACTCCCGGGTAAACCCACCG 61
 QY 212 TCAGAGCAGAGGTTTCTCTGATGACGGTTCCACGGTTCGCGAGTGTCCATCCGGCG 271
 DB 62 TCAGAGTTCGAGGTCGCACTCGACGACGGAGCGATTGCCCGGGCGGTGTCCCGAGTGGCG 121
 QY 272 CATCCACCGCGCTCCACGAGGCTCATGAGTCGCTGACGCTGCGGATTCGCTACCTGGGCA 331
 DB 122 CGTCCACCGGTCGATTGGAAGCGGTGAACTGCGCGATGGTGGACCGTTACCGGGGTA 181
 QY 332 AGGGCGTTTGAAGGAGTTGAAACGTCACAGGAAATCGGGGACGAGCTCGTGGCC 391
 DB 182 AGGGCGTGGAGAGGCGGTCCCGCGGTCAACGAGGAAATCTCCGAACAGATCGTGGGTT 241
 QY 392 TAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATGATGATGAGTGGACCGCCA 451
 DB 242 ACAGAGCCGAGGACGCGGTGTGACAGAGCGCCCTGATCACTTGGACGATGATCCCGC 301
 QY 452 ACAAGTCCCGCTGGGTGCAACGCAATCTCTGGTGTTCATGGCTGTGCAAGGCTG 511
 DB 302 ACAAGTCCCGCTGGGCGCCCAACGCAATCTCTGGTGTTCATGGCTGTGCAAGGCTG 511

QY 512 CTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACACGTTTC 571
 DB 362 CCGCGAGAGCGCGGACCTCGCGCTTTCGCTACTTGGCGGCGCCGAAACCCACGTCG 421
 QY 572 TTCAGTTTCCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCCGGTGTGACG 631
 DB 422 TGGCGGTGCCGATGATGAACATCTCAACGGGTGGCGACGCGGACCAACGTCGACA 481
 QY 632 TTCAAGNATTCATGATCGCTCCATCGGTGGAGAGACTTCTCTGAGGCTTCTCGCAACG 691
 DB 482 TCAGAGGATTCATGATCGCCCTCCATCGGGGCGGAGAGCTTCCGGGAGGCGTGGCTGGG 541
 QY 692 GCGCGAGGCTTACACGCACTGAAGTCGCTCATCAAGGAAAGGGCTGTTCACACCGGAC 751
 DB 542 GCGCGAGGCTTACCACTCGCTCAAGGGGTGCTCAAGGCCACGCGGTGCGCCACCGGTG 601
 QY 752 TTGGCGATGAGGGCGGCTTCGCTCTCTTCCTTCGCTCCACCGCTGAGGCTCTTGAACCTTA 811
 DB 602 TCGGTGACGAGGGCGGCTTCGCGCCAACTGGACAGCAACCGCGCGCTCGACCTGA 661
 QY 812 TCGTTAAGGCAATCGAGAAGGCTGGCTTCAACCCAGGCAAGGACATCGCTTTCCTCTGG 871
 DB 662 TCAGCGAAGCCATCACGAAGGCGGCTTCACGCTGGGCGCAGGACATCGCCCTGCTTGG 721
 QY 872 ACGTTGCTTCTCTGAGTTCTTCAAGGACGCGCACTTACCACTTCGAAAGTGGCGCAGCACT 931
 DB 722 ACGTGGCGGACCGAGTTCTTACGCGACGCGGCTTACAGTTCGAGGGTTCAGGCGCA 781
 QY 932 CCGCAGCTGAGATGGCAACGTTTACGCTGAGCTGCTGTGACGGTACCCCAATCGTCTCCA 991
 DB 782 GCGCGGAGGAGATGGCGGCTTACTACACCGAGCTGGTTCGAGTCTTACCGCTGGTTCGA 841
 QY 992 TCAGAGACCCATCGCAGGAAGATGACTGGGAGGTTTACCAACCTCACCGCAACCACTCG 1051
 DB 842 TCAGAGACCCGCTCAGCGAGGAAGACTGGGCGGCTGGAAGGCGCTCACCGAGTCCCTCG 901
 QY 1052 GCCACAAGGTTTCAAGTTCGTTGGCGGCGACTTCTTCGTCACCAACCTGAGCGCTGAAGG 1111
 DB 902 GCGACCGGCTGCAACTGCTCGGCGACGACCTGTTTCGTCACCAACCGGAGCGGCTGAGC 961
 QY 1112 AGGGCATTCGTAAGAGGCTGCCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCC 1171
 DB 962 GCGGTATCGACGAGGGCGCTGCCAACTCGCTGCTGTTCAAGTGAACAGATCGGTACGC 1021
 QY 1172 TCACCGAGACTTCGAGCTGTGCAATGGCTCACCGCGCAGGCTACCTTCCATGATGT 1231
 DB 1022 TCAGCGAGACTTTCGAGCGGCTGTGCTGGCGAGCGCAACCGGTACACCGGATGATCA 1081
 QY 1232 CCACCGTTCCCGGTGAGCCGAGGACACCACTTGTGCTGACCTCGGAGTTCGACTCAACT 1291
 DB 1082 GCCACCGGTTCGTTGAGACCGGAGACACCACTGCGCGGATCGCGGTGGCGGCAACT 1141
 QY 1292 GTGGCCAGATCAAGACTGCTGCTCCAGCAGCTTCCGACCGTGTTCGCAAGGTACCAACGAGC 1351
 DB 1142 CGGGGAGATCAAGACGGGTGCCCGCGCGCAGGAGCGGTCGCCAAGTACCAACGAGC 1201
 QY 1352 TTCTCCGATTCGAGAGCTGCTTGGGCAACGCGCGGCTGTACGAGGTTTCGAGCGCATTC 1411
 DB 1202 TGTGCGGATTCGAAGAAGAGCTCGACGACGAGCGGCTTACGCGCGGAGCAACGCGTTTC 1261
 QY 1412 CACGCTT 1418
 DB 1262 CCGCTT 1268

RESULT 12
 AC38079
 ID AC38079 standard; DNA; 1287 BP.
 XX
 AC38079;
 AC
 XX
 DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #19736.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW drug design; gene.

XX OS Mycobacterium avium.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362659P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR P-PSDB; ABU34209.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 14; SEQ ID NO 25949; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

XX CC proliferation; (7) identifying a compound that influences the activity of

XX CC the gene product or that has an activity against a biological pathway

XX CC required for proliferation, or that inhibits cellular proliferation; (8)

XX CC identifying a gene required for cellular proliferation or the biological

XX CC pathway in which a proliferation-required gene or its gene product lies

XX CC or a gene on which the test compound that inhibits proliferation of an

XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX CC compound's activity; (11) a culture comprising strains in which the gene

XX CC product is overexpressed or underexpressed; (12) determining the extent

XX CC to which each of the strains is present in a culture or collection of

XX CC strains; or (13) identifying the target of a compound that inhibits the

XX CC proliferation of an organism. The antisense nucleic acids are useful for

XX CC identifying proteins or screening for homologous nucleic acids required

XX CC for cellular proliferation to isolate candidate molecules for rational

XX CC drug discovery programs, or for screening homologous nucleic acids

XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

XX CC prokaryotic essential genes. Note: The sequence data for this patent did

XX CC not form part of the printed specification, but was obtained in

XX CC electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1287 BP; 214 A; 424 C; 475 G; 174 T; 0 U; 0 Other;

Query Match 41.08; Score 647; DB 8; Length 1287;

Best Local Similarity 69.98; Pred. No. 1.2e-153;

Matches 889; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

151 GTGGCTGAATCATGCACGATTGCTCGCGGAAATTCGACTCCCGGGTAAACCAACC 210

1 GTCCCGATTATCGAGCAGGTCTGGGGCCCGCGAGATCTCGACTCCCGCGGTAAACCCGACA 60

211 GTCGAGCGCAGAGGTTTTCTTGGATGACGGTTCCACCGTGTCCAGGTGTTCATCCGGC 270

61 GTCGAGGTCGAGATCGCCCTGACCGACGGAACGTTGCGCCCGCGCGGGTGTCCGTCGGG 120

271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACCGTGGCGATCGTACTCTGGC 330

121 GCCTGACCGGTCGAGCAGCAGGCTGCGAACTCGCGACCGCGGGGAGCGGTACGGCGGC 180

331 AAGGCGGTTTTGAAGGCAGTTGAAAAAGCTCAAAGAAATCGCGACGACGAGTCGCTGGC 390

181 AAGGGGTGCGAAGAGGGGTGCGAGGCGTCTCGACGAGATCGCCCGCGGTGATCGGG 240

391 CTAGAGGCTGACGATCAGCGCCCTCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGCC 450

241 CTGAACGCGCAGCAGCAGCGCTGGTGGACACGAGCGCTCGTGCACCTGGACGCGCACCCCG 300

451 ACAAAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATGGCTGTTCGAAAGGCT 510

301 GACAAGTCTGAGGCTGGCGGCAACGCGATCTGGGGGTGTCTGGCGGTGGGCGAAGGCG 360

511 GCTGCTGATTCCGACAGGCTCCCACTGTTCCGCTATCATCGGTGGACCAACGACACAGCTT 570

361 GCGCGGATTCGCGGAGTTGCGGCTGTTCCGCTACTCTGGCGGGCCCAAGCGCACATC 420

571 CTTCCAGTTCCAAATGATGAACATCATCAACGGTGGGCTCACGCTGACTCGGTGTTGAC 630

421 CTGCGGCTGCGATGATGAACATCTCAACGGGCTGGCGACGCGACACCGCGCTCGAC 480

631 GTTCAGGAATTCATGATCGCTCCAATCGGTGAGAGACCTTCTTGAGGCTCTCCGCAAC 690

481 ATCCAGAGTTTCATGTTGCGCCCGATCGGCGCGCGAGTTTGCCTGAGGATTCGCGGTGG 540

691 GCGCGCGAGGTTCTACCAACGACCTGAAAGTCCGTTCATCAAGGAAAGAGGCTCTCACCGGA 750

541 GGTGCGGAGGTGTATCACTCGCTGAGTCCGTGCTGAAGAGAGAGGCTCTGAGCACCGGC 600

751 CTTGGCGATGAGGGCGGCTTCGCTCTTCGCTCGGCTCCACCGTGGAGGCTCTTGACCTT 810

601 CTGGCGGACGAGGGCGGCTTCGCCCCCGAGCTGGCGGCGCACACCGCGGCGCTGGACCTG 660

811 ATCGTTAAGGCAATCCGAGAGGCTGGCTTCACCCCGAGGAGGACATCGCTCTTGCTCTG 870

661 ATCGCGCGGCGCATCGAATCGGCGGCTTCAAATCGGACCGGCTGAGGCTGGGCGCTG 720

871 GAGCTTGTCTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACCTTCGAAGGTGGCGAG 927

721 GACGCGCGCGCCACCGAGTTCTACAGCGACGGCACCGGCTACAAGTTTCGAGGGCGAGCAC 780

928 CATCTCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTCGTTGACCGGTACCCCAATCGTC 987

781 CGCACCGCGGAGCAGATGGCGGAGTTCTACGCGGGGCTGCTCGCGCGGTATCCGTTGGTG 840

988 TCCATCGAGGACCCACTGACGAGGAGATGACTGGGAGGTTTACACCACTTCACCGCAACC 1047

841 TCCATCGAAGACCCGCTGTCGGAAGATGATGGACCGCTGGGCGGCGCTGACCGCGCTG 900

1048 ATCGGGGACAAAGTTTCAGATCGTTGGCGAGACTTCTTCGTCACCAACCTTCGAGCGGCTG 1107

901 ATCGGGGACCGGGTGCAGCTCGTCGCGGACGACGCTCTTCGTCAGCAACCCCGAACGCGCTG 960

1108 AAGAGGGGATCGCTAAGAGGCTGCCAATCTCATCTCTGGTTAAGGTGAACCATCGGT 1167

961 GAAAGAGGGTATCGAAGAGGGCGTCGCAAAATGCGTTGCTGCTGAAGGTGAATCATGATCGGC 1020

1168 ACCCTCACCGAGACCTTCGACGCTGTGACATGGCTCACCGCGAGGCTACACCTCCATG 1227

1021 ACCTGACCGAGAGCTGGACGCGCTGCGCTGGCCCAACACAGCGGCTACCGCACGATG 1080

1228 ATGTCCACCGCTTCGGTGAGACCGAGGACCAACCATTCGCTGACCTTCGAGTTGCACTC 1287

Db 1081 ATGAGCCACCCAGCGGTGAACCCAGAGGACACACAGATCGCCGACCTGGCGGTGGCCGTC 1140
Qy 1288 AACTGTGCCAGATCAAGACTGTGTGCTCCAGCAGCTTCCGACCGGTGCGAAAGTACAAC 1347
Db 1141 GGACGCGGAGATCAAGACCGAGCGCGGCCGCGAGCGGTGGCCAACTACAC 1200
Qy 1348 CAGCTTCTCGCATCGAGCAGCTGCTGTGGCGACGCCGCGGTCTACGAGGTGCGAGCGCA 1407
Db 1201 CAGCTGTGCGGATCGAGGAGCGCTCGGCACGCGCGCTTACGCGCGCGACCTGGCC 1260
Qy 1408 TTCCACGCTT 1418
Db 1261 TTCCCGCGTT 1271
RESULT 13
AAS59564/c
ID AAS59564 standard; DNA; 30159 BP.
XX AAS59564;
XX DT
XX 13-FEB-2002 (first entry)
XX Propionibacterium acnes immunogenic protein encoding DNA #59.
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.
XX Propionibacterium acnes.
OS WO200181581-A2.
XX PN
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX Claim 1; SEQ ID NO 59; 1069pp; English.
XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by P. acnes. The disorders include
CC osteomyelitis (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of P. acnes in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for P. acnes proteins. These antibodies can be
CC used to downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU54183-AAU54498 and AAU67555-AAU67557. Note: The

CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 30159 BP; 6067 A; 9304 C; 8907 G; 5878 T; 0 U; 3 Other;
Query Match 40.6%; Score 640.8; DB 4; Length 30159;
Best Local Similarity 69.2%; Pred. No. 1.3e-151;
Matches 889; Conservative 0; Mismatches 392; Indels 3; Gaps 1;
Qy 152 TGGCTGAAATCATGCACGTATTCCGCTCGGAAATTCGACTCCGCGGTGAACCAACCG 211
Db 1539 TGGCAACCATCGAATTTCAGAGCCCGTGAGATTCCTCGATTCCCGCGCAACCGACCG 1480
Qy 212 TCGAGGCGAGAGGTTTTCTCGATGACGGTTCACACGGTTCGACAGGTTCATCTCCGCGC 271
Db 1479 TTGAGGTGAGATGATCTCGACGACGCGCACCCAGGCCCGCTGCGGTTCTTGGGGCG 1420
Qy 272 CATCCACCGCGTCCACGAGGCTCATGAGCTGCGTGACGG---TGGCGATCGCTACCTGG 328
Db 1419 CTTTCGACCGGTCAAGTTCGAGGCGGTTCGAGCTGCGTGACGGCGATAAGAAGCGTTACTCCG 1360
Qy 329 GCAAGGCGGTTTTGAAGGCGAGTTGAAACCGTCAACGAAAGAAATCGGCGACGAGCTCGTG 388
Db 1359 GTAAGGGTGTCTCAAGGCTGTTGAGAACTGTCACGAGAAAGATCGCCGAGGAGGTGCTCG 1300
Qy 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGTATGCGACCG 448
Db 1299 GCTGTGAGCGAGTGACGAGCGCATTTATCGACAGATCATGATCGAGCTCGACGATCCG 1240
Qy 449 CCAACAGTCCCGCTGGGTGCAACGCAATCTTTGGTGTTCATGGCTGTTCGCAAGG 508
Db 1239 ACAACAGGCGCAAGTTGGGTGCTAAACGCCATCTTTGGTGTCTCCCTGGCTGCTCATG 1180
Qy 509 CTGCTGCTGATTCGCGAGGCTTCCACTGTTCGCTACATCGGTGAGCAACGACACG 568
Db 1179 CTGCTGCTGATGTGCGGAGCTTCCCGCTGTACCAAGTACCTCGGCGGACCGAACTCCACG 1120
Qy 569 TTCTTCCAGTTCCAATGATGAACATCATCAACGTCGCGCTCAGCTGACTCCCGGTGTTG 628
Db 1119 TGCTGCCGTTCCAATGATGAACATCTCAACGCTGCTGCCACGCGGATTTCCGACGTTG 1060
Qy 629 ACCTTCAGGAATTCATGATCGCTCCAAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGCA 688
Db 1059 ACATCAGGAGTTTCATGATTCGCCCGATCGCGCGGAATCTTTAAGCAGGCTACAGGT 1000
Qy 689 ACGGCGGAGGCTTACCACGCACTGAAGTCGCTATCAAGGAAAGGGGCTGTGCACCG 748
Db 999 GGGGCGCTGCGGTTTACCACTCCCTCAAGAAAGGTTCTCAAGGACACAGGGCTTGGCTACCG 940
Qy 749 GACTTGGCGATGAGGGGCGCTTCGCTCTTCCTCGGTCCGCTCCACCGTGAGGCTCTTGACC 808
Db 939 GTCTGGGTGACGAGGGCGGTTTCGCCCCCAACCTTCCACAGAACGCCGCGCTTTGACC 880
Qy 809 TTATCGTTAAGGCAATCGAGAAGGCTGGCTTCACCCCGAGGAGACATCGCTTTGCTC 868
Db 879 TCATCTTGAAGCATCAAGCGCGCGCTTCGAGCGGGTGAAGACGTGCGCTTCGCC 820
Qy 869 TGAAGGTTGCTTCTCTGATTTCTTCAAGGACGCGCACTTACCACTTCAAGAGTGGCGCAGC 928
Db 819 TTGATGTGCTGCTTCCGAATTTCTTTCGAGGACGCGCAAGTACAGTTTCGAGGGCGCAGGCTA 760
Qy 929 ACTTCCGCGCTGAGATGGCAACGTTTACCTGAGCTCGTTGAGCGCTTACCAATCTCT 988
Db 759 AGACCTCGGCTGAGATGATCGCGTATCTACGAAGCGCTTCATCGGCAAGTATCCGCTGCT 700
Qy 989 CCATCAGGAGCCCACTGCAGGAAGATGATCTGGAGGGGTTACACCAACCTCACCGCAACCA 1048
Db 699 CTATTGAGGATCCGTTGGACGAGGAGATTGGACGCGGTGGGCGGAGTTTCAAGAGAGC 640
Qy 1049 TCGGCGACAGGTTTCAAGATCGTTGGCGACGACTTCTTTCGTCAACCAACCTTGAGCGCTGA 1108
Db 639 TTGGTGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCACTTAACCCGAGGCGCTTG 580

Db 819 TTGATGTGCTGCTTCGGAATTTTCGAGGACGCAAGTACACGTTTCGAGGGCCAGGCTA 760
 Qy 929 ACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCTACCCAACTCCTCT 988
 Db 759 AGACCTCGGCTGAGATCGCGTACTACGAAGGCTCATCGCCAAATATCCGCTGCTCT 700
 Qy 989 CCATCGAGGACCACTCGAGAGATGACTGGGAGGTTTACCAACATCAGCGCAACCA 1048
 Db 699 CTATTGAGGATCCGTTGGACGAGGAGATTGGGACGCTTGGCCGAGTTTACGAAGAAGC 640
 Qy 1049 TCSCGACAGAGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAACCTGAGGCGCTGA 1108
 Db 639 TTGTGAGAAGATCCAGATCGTTGGCGACGACTTGTTCGTCACTAACCCGAAGCGCCTTG 580
 Qy 1109 AGGAGGCGATCGCTTAAGAAGGCTGCCAACTCCATCTCGTTAAGGTGAACAGATCGGTA 1168
 Db 579 CCAAGGGAATCGAGACCAAGGCTGCCAAAGCCCTGCTCGTCAAGGTGAACAGATTGGCT 520
 Qy 1169 CCTCACCAGACTTCGAGCGTGTGACATGCTGCTACCGCGCAGGCTACACCTCCATGA 1228
 Db 519 CCTCTCGGAGACTATCGAGCGCTCGAGCTGGGCCCATCGCAACGGCTACCGCTGCATGA 460
 Qy 1229 TGTCCCAACCTTCGCTGAGACGAGGACACCACTTGTGCTGACCTCGAGTTGCATCA 1288
 Db 459 TGTGCAACCGTTCTGGTGAGACCGAGGACCACTATCGCGGATCTCGCCGCTCGCGCTAT 400
 Qy 1289 ACTGTGGCCAGATCAAGACTTGGTCTCCAGCAGCTTCCGACCGTTCGCAAGTACAAAC 1348
 Db 399 CGACTGTGAGATTAAATCCGTTGCCCGCCCGCGGAGCGGATCGCCAGTACAAAC 340
 Qy 1349 AGCTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGCTTACGCGAGGTCGACGCGCAT 1408
 Db 339 AGTTGCTGCTATTGAGGAGAACTGGGCGACTCGGCGGAGTACGCGGCTGCTTCCGCTT 280
 Qy 1409 TCCACGCTTTCAGGGCTAATAA 1432
 Db 279 TCCGCGCTTCCAGGCGCTGACTGA 256

RESULT 15

ABX09144
 ID ABX09144 standard; DNA; 45190 BP.

AC AC
 XX XX

DT 08-APR-2003 (first entry)

XX Mycobacterium tuberculosis H37Rv BAC clone BAC-Rv267.

DE Mycobacterioses; survival; virulence; protective antigen; vaccine;
 KW mycobacterial disease; tuberculosis; leprosy; ds; cosmid.

XX Mycobacterium tuberculosis.

XX WO200274903-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-IB001973.

XX 22-FEB-2001; 2001US-0270123P.

XX (INSP) INST PASTEUR.

XX Cole S;

XX WPI; 2002-759885/82.

XX Identifying and selecting genes for survival or virulence of mycobacteria
 PT by a comparative genomic analysis of the sequences of Mycobacterium
 PT tuberculosis and M. lepreae.

XX Disclosure; Fig 8; 874pp; English.

XX

CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC lepreae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a Mycobacterial cosmid DNA
 CC sequence used in the method of the invention

SQ Sequence 45190 BP; 7490 A; 14522 C; 15204 G; 7974 T; 0 U; 0 Other;

Query Match 40.1%; Score 632.6; DB 6; Length 45190;
 Best Local Similarity 68.4%; Pred. No. 1.8e-149;
 Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGCTGAATCATGCACCTATTTCGCTCGGAAATTCGACTCCCGCG 199
 Db 19931 AGAGAAACCCAGTGCCTGATTCGAGCAGGTTAGGCCCCGAGAGATCCTCGATTCCCGCG 19990
 Qy 200 GTAACCCACACCGTCCGAGGAGAGTTTCTCGATGACGGTTCACCGGTTCGACAGTG 259
 Db 19991 GCACCCGAGGTTGAGGTCGAGTGGCGCTTATCGACGGGACATTCGCCGGGCCCGCG 20050
 Qy 260 TTCCATCCCGCGCATCCACCGCGCTCCACGAGCTCATGAGCTGCGTGCAGCGTGGCGATC 319
 Db 20051 TGCCTGTCGGCGCTCGACCGGGGAGCAGGAGCGTTCGAGTTGCGCGACGCGCGCATC 20110
 Qy 320 GCTACTCTGGCAAGGGCGTTTGAAGCAGTTGAAACGTCACAGAGAAATCCGCGACG 379
 Db 20111 GCTACCGCGCAAGGCGTGCACAAAGCCGTCGAGGCTGTTCTTGATGAGATCGGCCGG 20170
 Qy 380 AGCTCGCTGGCTAGAGGCTGACGATCAGCGCTCATCGACAGCAATGATCAAGCTTG 439
 Db 20171 CCGTATCGGACTCAACCGCGACGACGAGGATTTGTCGACAGGCGCTGGTGACCTAG 20230
 Qy 440 ATGGCACCGCCAAAGTCCCGCTCGGTGTCACAAACGCAATCCTTGGTGTTCATGCTG 499
 Db 20231 ACGGCACCCCGACAAAGTCCCGCTCGGTGTCACAAACGCAATCCTTGGTGTTCATGCTG 20290
 Qy 500 TTGCAAGGCTGCTGCTGATTCGCGAGGCTCCACCTGTCGCTACATCGTGGACCAA 559
 Db 20291 TTGCCAAGCGCGCGGATTCGCGGAGCTGCGCTTGTTCGTTATGTCGGGGGGCCAA 20350
 Qy 560 ACGCACACGTTCTTCAGTTCCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACT 619
 Db 20351 ACGCGCACATTCTGCGGTACCGATGATGAACATCTCAACGGCGGCGCACACGCCGATA 20410
 Qy 620 CCGGTGTGACGTTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGAGACCTTCTCTGAGG 679
 Db 20411 CCGCTGTGACATTCAAGAGTTTCAATGGCGCAATTGGCGCGCCAGCTTCTGTCAGG 20470
 Qy 680 CTCTCCGCAACGGCGCGGAGTCTACACGCACTGAAGTCCGTCATCAAGGAAAAGGGCC 739
 Db 20471 CGTTGCGCTGGGGCGCTGAGGTGTACACGCGCTCAAGTCCGCTCTGAAAAAGAGGGGC 20530
 Qy 740 TGTCCACCGGACTTGGCGATGAGGGCGGCTTCCGCTCCTTCCGTCGGCTCCACCGCTGAGG 799
 Db 20531 TGTCCACCGGCTGGGGGAGCAAGGGGCTTCCGCCCGGATGTGGCGCGGACCAACCGCGG 20590
 Qy 800 CTCTTGACCTTATCGTTAAGGCAATCGAAGAGGCTGCTTACCCCGGCAAGGACATCG 859
 Db 20591 CGTTGACCTGATCAGCGGGGCCATCGAGTCCGGGCTTTCGACCGCGCGCGGACGCTGG 20650
 Qy 860 CTCTTGCTTGGAGCTTGTCTCTCTGAGTTCTTCAAGGACGGCAC---CTACACATTCG 916
 Db 20651 CGCTGGCCCTGGAGCGCGCGGCCACCGAGTTCTTACCCGACGCGCACCGGCTACGCTTCG 20710

Qy	917	AAGTGGCCAGCACTCGCAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGGCT	976
Db	20711	AGGGCAACCCGTACCGCAGACGAGATGACCGAGTTCTACGCGGGCTGCTCGGCGCCT	20770
Qy	977	ACCAATCGTCTCATTCGAGGACCCACTCGAGGAAGATGACTGGGAGGGTTACACCAACC	1036
Db	20771	ACCGCTGGTGTGATCGAAGACCCACTGTCCGAGACGATTGGGACGGCTGGGCGCGC	20830
Qy	1037	TCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTCAACCAACC	1096
Db	20831	TGACGGCCTCGATCGGTGACCGGGTGCATAATCGTCGGCGACGACATCTTTGTCAACCAATC	20890
Qy	1097	CTGAGCGCTGAAGGAGGCGATCGCTAAGAGGCTGCCAACTCCATCCCTGGTTAAGGTGA	1156
Db	20891	CCGAGCGGCTCGAGGAGGGCATCGAACGCGGGCGTGGCAATGCGTTGCTGGTCAAGGTGA	20950
Qy	1157	ACCAGATCGGTACCTCACCGAGACCTTTCGACGCTGTTCGACATGGGTCAACGCGCAGGCT	1216
Db	20951	ACCAGATCGGACGTTGACCGAGACACTCGACGCGGTTCACGCTGGCTCACCCACGGCGGAT	21010
Qy	1217	ACACCTCCATGATGTCACCGGTTCCGTTCCGTTGAGACCGGAGGACACCAATTTGCTGACCTCG	1276
Db	21011	ACCGCACGATGATCAGTCACCGCAGTGGCGGACGCGAGGACACCATGATCGCCGACCTCG	21070
Qy	1277	CAGTTGCACTCAACTGTGGCCAGATCAAGACTGTTGCTCCAGCACGTTCCGACCGCTGTCG	1336
Db	21071	CGGTGGCCATCGGACGGGCGAGATCAAGACGGGCGGCTGCTCGCAGTGAGCGGCTCG	21130
Qy	1337	CAAAAGTACAAACCACTTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGGCTTACGCGAG	1396
Db	21131	CAAAATACACCACTGCTGGGATCGAAGAGGCGCTTGGGACGCGGCGCGCTACGCGG	21190
Qy	1397	GTCGACGCGATTCCCAACGTTTCAGGGCTAAATAAAGCGCT	1439
Db	21191	GCGACCTGGCATTTCTCTCGGTTGCGGTGCGAGACGAAATAGGT	21233

Search completed: September 28, 2005, 17:01:34
 Job time : 796.5 secs


```

Db      90  ACCCCACGCTGAGGTTGACCTGACACACGAGAGGCGCTCTTCGG---CTGGCGGTCC 146
Qy      263  CATCCGGCGCATCACCGGCGTCCACGAGGCTCATGAGCTGCGTGAAGTGGCGA---TC 319
Db      147  CCTCTGGTGCCTCGACTGGCATCTATGAGGCTGCGAGCTCCGTGACGGCGACAAGTCGC 206
Qy      320  GCTACCTGGCAAGGGGTTTTGAAGGAGTTGAAGAGCTGAAGAGTCAAGAAATCGGCGAG 379
Db      207  GCTACCTCGCAAGGGGTGCTGAAGGCGCTGAGAAAGTCAAGAGTCTCGGCGCCA 266
Qy      380  AGCTCGTGCCTAGAGGCTGACGATCAGCGCTCTCATCGACGAAGCAATGATCAAGCTTG 439
Db      267  AGCTGATCGGCTCGAGCTCACAGAGGCGGAGATCGAGAGGCTCATGCTCAGATCG 326
Qy      440  ATGGCACCGCAACAAGTCCCGCTGGGTGCAACGCAATCCCTTGGTGTTCATGCGGTG 499
Db      327  ACGGCACTGAGAAACAGACCCACTGGGCGCCAAAGCCATCTCTCGGCTGCTCATGTCGG 386
Qy      500  TTGCAAGGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGT----- 552
Db      387  TCTGCGCTGGCGCGCGCTTTCAGGGGCGCTGCCCTCTTACCCTACATCGCGAGCTGT 446
Qy      553  --GGACCAACGACACGTTCTTCCAGTTCCAAATGATGAACATCATCAACGCTGGCGCTC 610
Db      447  CGGGCAACAAGTCCGCCATCTGCTCTGCCCTCTTCAAGCTCATCAACGCGGTGAGC 506
Qy      611  ACGTGACTCCGGTGTGAAGTTAGAAATTCATGATCGCTCCAAATCGGTGCGAGACCT 670
Db      507  ACGGGGCAACAAGCTCGCCATGAGGAGTTTCATGATCTGCCCCACCGGGCGCACCTCGT 566
Qy      671  TCTCTGAGGCTCTCGCAACGGCGCGAGGTCTACCAAGGACACTGAAGTCCGTATCAAGG 730
Db      567  TCCACGAGGCGCTCGCATGCGCCGCGAGAGCTACCAACAACCTGAAGCTCGTATCAAGA 626
Qy      731  AAAAGGCG-----CTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTTCGTCG 784
Db      627  AGAGTAGCATGAGCGCGACCAACGTCGGCGAGGAGGTGGCTTCGCCCCCAACATCC 686
Qy      785  GCTCCACCGGTGAGGCTCTTGAACTTATCGTTAAGGCAATCGAGAAGGCTGGTTCAACC 844
Db      687  AGGCCAACCAAGAGGCGCTGAGCTCATCGTTGAGGCGCATCAAGCAGGCGCGGTACACTG 746
Qy      845  CAGCGAGGACATCGCTCTTCTCTCGACGCTGCTTCTCTGAGTTCTTCAAGAGCGCA 904
Db      747  GCAAGATCGAGATCGGCAATGAGCTGCGCGCTCGTCTGTTGAGAGCCCAAGAGTCGA 806
Qy      905  CCTACCACTTCGAAGGTGGCCAGCACTCCGAGCTGAGA----- 943
Db      807  AGTAGCACTCGCTTCAAGGTCCCGCGACAGAGACCCCGACATGCTGCTCTCGG 866
Qy      944  -----TGGCAACGTTTACGCTGAGCTGTTGAGCGGTACCCAAATGCTCTCCATCG 994
Db      867  GCGAGGCGCTGATCAAGCTCTACGAGGAGTGGACCTCGAAGTACCCGATCTGGTCGATCG 926
Qy      995  AGGACCACTCGAGGAGATCACTGGGAGGTTACCAACCTCACCGCAACCATCGCG 1054
Db      927  AGGACCCCTTCGACGAGCACTGGGCGCACTTACACCAAGGTTTACCAGCTGATTCGCA 986
Qy      1055  ACAAGGTTTCAATGTTGGCGACGACTTCTTCTGTCACCAACCTCGAGCGCTGAAGGAGG 1114
Db      987  ACAGATCCAGATCGTTCGGCGACGACCTGCTGCTGTCACCAACCCCAAGCGCATCGTAGG 1046
Qy      1115  GCATCGCTAAGAGGCTCGCAACTCCATCTCTGTTAAGGTGAACAGATCGGTACCTCA 1174
Db      1047  CCGCAACAAGAGGCGCTGCAACGCGCTGCTGCTCAAGCTCAAACAGATCGGCTCGGTCA 1106
Qy      1175  CCGAGACCTTCGAGCTCTCGACATGGCTCACCGCGAGGCTACACCTCCATGATGTCCT 1234
Db      1107  GCGAGGCGCTCGAGGCGTGGCGCTGGCGCGGAGGTCAACTGGGGCGTCAATGTTGTCG 1166
Qy      1235  ACCGTTCCGTTGAGACCGAGGACACCACTTGTGACCTCGAGTGTGCACTCAACTGTG 1294

```

```

Db      1167  ACCGCTCGGGCGAGACTGAGGACCGGTTTCATCGCCGCACTCGTCTCGGTCTGGGTGCG 1226
Qy      1295  GCCAGATCAAGACTGGTCTCCAGACGTTCCGACCGTGTGCGAAGTACACACGAGCTTC 1354
Db      1227  GCCAGATCAAGACTGGTGCCTCCCTCGCTCGAGGCGCTCGCAAGTACACACGAGCTCC 1286
Qy      1355  TCCGATCGAGCAGCTGTTGGCGACGCGCGGCTCTACGCG 1396
Db      1287  TCCGATCGAGGAGGCTCGGCGCAAGCCCACTACGCG 1328

RESULT 2
BH770842/c
LOCUS      BH770842
DEFINITION  LLMtag581 MG1363 Random Sequence Tag Library Lactococcus lactis
              subsp. cremoris genomic, genomic survey sequence.
ACCESSION  BH770842
VERSION    BH770842.1 GI:20373799
KEYWORDS   GSS.
SOURCE     Lactococcus lactis subsp. cremoris
ORGANISM   Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococaceae;
            Lactococcus.
REFERENCE  1 (bases 1 to 2491)
AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL   Sci. Aliments (2002) In press
COMMENT   Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            Best homologue in strain IL1403 is enoA (100%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 2463.
FEATURES   Location/Qualifiers
            source          1..2491
                        /organism="Lactococcus lactis subsp. cremoris"
                        /mol_type="genomic DNA"
                        /strain="MG1363"
                        /sub_species="cremoris"
                        /db_xref="taxon:1359"
                        /clone_lib="MG1363 Random Sequence Tag Library"
                        /notes="Vector: pSGMU2; Site 1: Small; Library of
                        chromosomal fragments of L.lactis strain MG1363 was
                        prepared by partial AluI digestion or by sonication."
ORIGIN
Query Match      25.3%; Score 399.2; DB 8; Length 2491;
Best Local Similarity 60.5%; Pred. No. 1.7e-96;
Matches 722; Conservative 0; Mismatches 448; Indels 24; Gaps 3;
Qy      259  GTTCCATCCGGCGATCCACGGCGCTCCACGAGGCTCATGAGCTGCGTGCAGTGGCGA- 317
Db      2485  GTACCTTCAGTCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2426
Qy      318  --TCGCTACCTGGGAGGCGGCTTTGAGGAGGTTGAAACCTCAACGAAGAATCGGC 375
Db      2425  TCTCGCTACACGAGCTTGGTACTCAAAAAGCTGTTGACAACTGAACCAATCATCGCT 2366
Qy      376  GAGGAGCTCGCTGGCTGAGGCTGAGGATCAGCGCTCATCGACGAAGCAATGATCAAG 435
Db      2365  GAAGCTATCATCGCTTATGAGTATCTGACCAACAGCTATTGACCGTGCAATGATCGCT 2306
Qy      436  CTTGATGGCACCGCCAAAGTCCCGCTCGGTGCAAAACGAAATCTTGTGTTTCATG 495
Db      2305  CTTGACGGTACTGAAAACAAAGGTAAATGGGAGCTAACGCTATTCTTGTGTTTCATC 2246
Qy      496  GCTGTTCAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGA 555

```

2245	Db	GCTGCTGCTGCTGCTGCTGATGAACATTTGGTGTTCACATTTTACAACCTACCTTTGGCGGA	2186
556	Qy	CCAAACGGACACAGTTTCTTCCAGTTTCCCAATGATGAACATCATCAACGGTGGCGCTCAGCGT	615
2185	Db	TTCAACGCTAAAGTATTTGGCCAACTCCCAATGATGAACATCATCAATGGTGGTCTTCTACTCA	2126
616	Qy	GACTCCGGTGTTCAGCTTTCAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCT	675
2125	Db	GAGCCCTTATCGCTTTTCCAAGAAATTCATGATCGTACCGTGGTGACCTCAATTCACAA	2066
676	Qy	GAGGCTCTCCGAACGGCGGGAGGTCTACCAACGACTGAAGTCCGTCATCAAGGAAAAG	735
2065	Db	GAAGCGCTTGGTGGGGTGTGTAATCTTCCACGCTCTTAAGAAAATCTTAAAGCTCGT	2006
736	Qy	GGCGTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTCTTCGTCGGCTCCACCCGT	795
2005	Db	GGACTTCAAAACAGCTGTCGGTGAAGGATTCGCTCCTTAAATTCGACGGAACGTAA	1946
796	Qy	GAGGCTCTTGACCTTATCGTTTAAGGCAATCGAGNAGGCTGCTTCACCCCGAGGCAAGGAC	855
1945	Db	GACGGTGTAGAAACTATCTTTAAAGCAATCGAAGCAGCTGGTTACAAGCTGGTGAGAT	1886
856	Qy	ATGCT-- -CTGCTCTGGAGTTGCTTCTCTGAGTCTTCAAGGACGCACTACCAC	912
1885	Db	GGCGTTATGATCGGTTTCGACTGTGCATCATCAGAAATCTACGAAAACGGTGTTCAGC	1826
913	Qy	TTCT-----GAAGTGGCGCAGCACCTCCGAGCTGAGATGGCAACGTT	954
1825	Db	TACACTAAATTCGAAGGTGAAGCGGTAAAAAATTTTCAGTCTGAACAAGTTGACTAC	1766
955	Qy	TACGCTGAGCTCGTTGAGCGGTACCCAATCGTCTCCATCGAGGACCCACTGCGAGGAAGAT	1014
1765	Db	CTTGAAGAACTCGTTTCTAAATACCCAATCATCACTATTGAAGATGTTATGCGCGAAAAC	1706
1015	Qy	GACTGGGAGGTTACACCAACTCACCGCAACCATCGCGACAGGTTACAGATCGTTGGC	1074
1705	Db	GACTGGGATGGATGGAAAAATCCTTACTGAAACGCTTGGTAAAAAGTTTCAACTCGTTGGT	1646
1075	Qy	GAGCACTTCTTCGTCAACCAACCTTGAGCGCTGAAGGAGGCGCATCGCTAAGAAGGCTGCC	1134
1645	Db	GAGCACTTCTTCGTACAAACACTAAATACCTTGAACGTGGTATCGTGAAATGCTTCA	1586
1135	Qy	AATCCCATCTGTTTAAGTGAACAGATCGGTACCTTCACCGAGACCTTTCGACGCTGC	1194
1585	Db	AACGCTATCTTGATCAAAGTTAACCAATCGGTACTTTTGACAGAAACTTTTGAAGCTATT	1526
1195	Qy	GACATGCTCACCGCGAGGCTACACCTCCATGATGTGCCACCGTTCGGGTGAGACCGAG	1254
1525	Db	GAATGGCTAAAGAAGCTGGTTTCACAGCAATCGTATCTCACCGCTTCAGGTGAACCTGAA	1466
1255	Qy	GACACCAACCATTCGTGACCTCGCAGTTTGCACCTCAACGTGGCCAGATCAAGACTGGTGCT	1314
1465	Db	GATTCACAAATCTCAGACATCGCTGTTTGCAACTNACGCTGGTCAAAATCAAACTGGTTCA	1406
1315	Qy	CCAGCAGCTTCGACCGGTGTCGAAAAGTACAACCAAGCTTCTCCGCAATCGCAGCTGCTT	1374
1405	Db	CTTTCAGGTACAGACCGTATGGCTTAATAACAACCAATTCGTTTGGTATTAAGACCAATTG	1346
1375	Qy	GGCGACCCGGGCTCTACGGCAGGTTCGCGCGCATTTCCCAACGCTTTCAGGCGTAA	1428
1345	Db	GCTGAAGTTGCTCAATACAAAGGCTTTAAAGCATTTCTACAACTTAAAAATAA	1292

RESULT 3

CR685174	CR685174	1626 bp	linear	HTC 19-AUG-2004
DEFINITION	Tetraodon nigroviridis	full-length	CDNA.	
ACCESSION	CR685174			
VERSION	CR685174.1			
KEYWORDS	HTC; CDNA; full-length; Tetraodon nigroviridis.			
SOURCE	Tetraodon nigroviridis			
ORGANISM	Tetraodon nigroviridis			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source	ORIGIN
1	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontoidea; Tetraodontidae; Tetraodon.	Genoscope.	Submitted (10-AUG-2004)	Genoscope - Centre National de Sequencage - 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) The sequences are based on single pass reads. More information available at http://www.genoscope.cns.fr/tetraodon.	Location/Qualifiers		
	1 (bases 1 to 1626)	Direct Submission			.1626		
					/organism="Tetraodon nigroviridis"		
					/mol_type="mRNA"		
					/db_xref="taxon:99883"		
					/tissue_type="Eyes"		
					Query Match 24.6%; Score 388.4; DB 3; Length 1626;		
					Best Local Similarity 61.0%; Pred. No. 1.3e-93;		
					Matches 799; Conservative 0; Mismatches 451; Indels 60; Gaps 8;		
Qy	160	ATCATGACGCTATTCGCTCCGGAATATTCGACTCCGCGGTAAACCAACGCTGAGGCA	219				
Db	110	ATCTGAAGATCACGCTCGGGAATATTCGACTCCGCGGCAACCCACCGCTGGAGTC	169				
Qy	220	GAGGTTTCTGATGACGCTCCCAAGGTTCGAGGTTCATCCGCGCATCCACC	279				
Db	170	GACCTGTACCAAGAAAGTCTGTTACAGAGTGGC	226				
Qy	280	GGCGTCCACGAGGCTCATGAGCTCGGTGACGGTGGCA	336				
Db	227	GGCATCTACGAGCCCTGAGCTCCGCAACGACAAACCCGCTACATGGGCAAGGT	286				
Qy	337	GTTTTGAAGCAGTTGAAAACGTTCAACGAAGAAATCGCGACGAGCTCGTGGCTAGAG	396				
Db	287	GTCTCTAAAGCTGTGAGAAATATCAATAAATTAATTCACCTGCTGGTTGGCAAGAC	346				
Qy	397	GCTGAGCATCAGGCC	450				
Db	347	GTCACGTTCTGACCAAGGCAAAATTCGACAACTGATGCTGGAGATGGACGCGCAC	406				
Qy	451	AACAAGTCCCGCTGGGTGCAACGCAATCCTTTGGTGTTCATGGCTGTGTCAAAGGCT	510				
Db	407	AACAATCTAAGTTTGGGGCAACGCCATCTCGGGGTGTCCTTGGCTGTGTGCAAGCT	466				
Qy	511	GCTGCTGATTCGACGCTCCCACTGTTCCGCTACATCGGTGGAACAAACGC	563				
Db	467	GGTGCAGCAGAGAAGGCGTGCACCTACCGCCACATCGCGACCTGGCGGCAACCCC	526				
Qy	564	--ACAGTTCCTTCAGTTTCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCC	621				
Db	527	CAAGTCATCTCTCCCGTTCCTCGCTTTCAACGTCATCAACGGCGGCTCCCAACGAGAAC	586				
Qy	622	GGTGTGACCTTCAGGAATTCATGATCGCTCCAAATCGGTGCAGAGACCTTCTCTGAGGCT	681				
Db	587	AAGCTGCCATCAGAGATTTCATGATCTCGCGGTGGAGCCAGACAGCTTCAAGAGGCC	646				
Qy	682	CTCCGCAACGGCGGAGGTCTACCAACGCACTGAAAGTCCGTCATCAAGGAAAAGGCC	738				
Db	647	ATCGCATCGGCGCGAAGTCTACCAACACTGAAGAACGTCATCAAGGAGAGTACGGC	706				
Qy	739	---CTGTCCACCGGACTTGGCGATGAGGGCGGCTTTCGCTCTTCGTCGCTCAACCGT	795				
Db	707	AAGGACGCCCAACAGCTGGGAGACGAGGGAGGCTTTGCCCCCAACATCTCTGGAGAACAG	766				
Qy	796	GAGGCTCTTGACCTTATCGTTAAGGAATTCGACGAAGCTGGCTTCAACCGACGAGGAC	855				
Db	767	GAGCTCTGAGCTGTGAAGAACGCCATCGTCAAGGCGGCTACAC	820				
Qy	856	ATCGCTCTTGACGCTGCTTCCTCTGAGTCTTCTCAAGGACGCGCACTTACCCTTC	915				

```

Db      821  ATTGTGATCGGATGGACGTGGCTGCTCTGAGTTCTACAAGGGCGGCAAGTACGACCTG 880
Qy      916  GAA-----GGTGGCAGACCTCCCGACGTGAGATGGCAACGTT 954
Db      881  GACTTCAGTCTCCGACGACCCGGCGGTATCATCTCTCCGAGCAGCTGGCGACCTC 940
Qy      955  TAGCTGAGTCTGTGACGGGTACCCAAATCGTTCATCGATGAGAACCCACTGCAAGAGAT 1014
Db      941  TACAGGGGCTTCTGCAAGATTACCCAGTGTGTCTCATCGAGACCCCTTTTGACAGGAC 1000
Qy      1015  GACTGGGAGGTTACACCAACCTCACGCAACCATCGGCGACAAAGTTCAGATCGTTGGC 1074
Db      1001  GACTGGGAGGCTGGACCAATTTACAGCCAGCACCAAGC-----ATTCAAGTGGTGGG 1054
Qy      1075  GACGACTTCTGCTGACCAACCTTACGCGCTGAAGGAGGGCATCGTAGAAGGCTGGC 1134
Db      1055  GACGACTTCAAGTACCGGACCCCAACGATCGCCAAAGGTGTGGCCAGAGTCTCTGC 1114
Qy      1135  AACTCCATCTGTTAAGGTGAACAGATCGGTACCTCTCCAGAGCTTTCGACGCTGTC 1194
Db      1115  AACTGCTGCTGCTCAAGTCAACAGATCGGTCTGGTCAACGAGTCCCTGCGAGCATGC 1174
Qy      1195  GACATGCTCACCGGAGGCTACCTCCATGATGTCACCGTTCGCGTGAGACGAG 1254
Db      1175  AAGATGGCCAGAGCAACGGCTGGGGCTGATGTCAGCCATCGCTCCGGGGAGACGGAG 1234
Qy      1255  GACACCAACCATGCTGACCTCGCAGTTGCACTCAACTGTGTGGCCAGATCAAGACTGTGCT 1314
Db      1235  GACACCTTCACTGCTGACCTGGTGTGGTCTGTGACCGGACAGATCAAGACGGGGCT 1294
Qy      1315  CCAGCAGCTTCGACCGCTGCGAAAGTACAACAGCTTCTCCGATCGAGCAGCTGCTT 1374
Db      1295  CCGTGGCTGCTCGAGCGCTGGCCAAAGTACAACAGCTGCTCCGATCGAGGAAGCTG 1354
Qy      1375  GCGACGCGCGCTGACGAGGTCGAGCGCATTCGCCAGCTTTCAAGG 1424
Db      1355  GCGACGAGGCGCTTCGCCGGCAAGAACTTCAGGCAACCCATCTGAGCG 1404

```

RESULT 4

```

LOCUS      CR726296      1643 bp      mRNA      linear      HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION  CR726296
VERSION    CR726296.1  GI:51244547
KEYWORDS   HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE     Tetraodon nigroviridis

```

ORGANISM

```

Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1643)

```

REFERENCE

```

AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
           : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
           (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT    The sequences are based on single pass reads.
           More information available at
           http://www.genoscope.cns.fr/tetraodon.

```

FEATURES

```

source
Location/Qualifiers
1..1643
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

```

ORIGIN

```

Query Match      24.5%; Score 386.8; DB 3; Length 1643;
Best Local Similarity 60.9%; Pred. No. 3.4e-93;
Matches 798; Conservative 0; Mismatches 452; Indels 60; Gaps 8;

```

```

Qy      160  ATCATGACGATTTCTGCTCGGAAATTTCTGACTCCCGGGTAAACCCACCGTCGAGCA 219
Db      113  ATCTTGAAGATCCACGCTCGGAAATATATTCGACTCCCGCGCAACCCACCGTGAGGTC 172
Qy      220  GAGGTTTTCTCGATGACGGTTCCACGCTGCGAGGTGTTCATTCGCGGCATCCACC 279
Db      173  GACCTGTACACAAAGAGTCTGTTACAGCTGGG---GTCCCAGCGCGCTCCACC 229
Qy      280  GCGCTCCAGAGGCTCATGAGCTGCGTGAAGGTGCGA---TCGTACTCTGGGCAAGGCG 336
Db      230  GGCATCTACAGGCTTGGAGCTCCGCGCAACGCAAAACCCGCTACATGCGCAAGGT 289
Qy      337  GTTTTGAAGGCAATGAAACGTCACGAGAAATCGGCGAGAGCTCGTGGCCCTAGAG 396
Db      290  GTCTCTAAAGCTGTGAGAAATCAATAAATTAATTCACCTGCACTGTTGGCAAGGAC 349
Qy      397  GCTGACGATCAGCGCC-----TCATCGACGAAAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db      350  GTCAAGCTTCTGGACCGCAAAATCGACAGCTGATGCTGGAGATGACGCGCACCGAC 409
Qy      451  AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTTCCATGGCTGTTGCAAGGCT 510
Db      410  AACAAATCTAAGTTCCGGGCAACGCAATCTTGGGCTGTCCCTGGCTGTGTGCAAGGCT 469
Qy      511  GCTGCTGATTCGCGAGGCTCCACCTGTTCCGCTACATCGGTGGACCAAAACGC-----563
Db      470  GGTGACGACGAGAAAGGCGTGGCCCTCTACCGCCATCCCGCATCTGGCGGCAACCC 529
Qy      564  --ACAGTTCTTCCAGTTCCAATGATGAACATCATCAACGGTGGGCTCACGCTGACTCC 621
Db      530  CAAAGTCACTCTCCCGTTCCGCTTTCAGCTCATCAACGGGCTCCACGCGAGCAAC 589
Qy      622  GGTGTTGAGCTTCAAGAAATTCATGATCGCTCCATCGGTGCGAGACCTTCTCTGAGGCT 681
Db      590  AAGCTGGCCATGACGAGTTTCATGATCTCTCCGCTGGGAGCCAGCAGCTTCAAAGAGGCC 649
Qy      682  CTCCGCAACGGCGGAGGCTTACCAACGCTGAGTCCGCTCATCAAGGAAAGAGGCC---738
Db      650  ATGCGCATCGGCGCCGAGTCTACCAACCTGAGAACGTCATCAAGAGAGATGACGGC 709
Qy      739  ---CTGTCACCGGACTTGGCGATGAGGGGCTTCGCTCTTCTCCGCTCCACCCGT 795
Db      710  AAGGATGCCCAACCAACGTTGGGAGACGAGGAGGCTTTTCCCGCAACATCTGGAGAAACAAG 769
Qy      796  GAGGCTCTTGAACCTTATCGTTAAGGCAATCGAGAGGCTGGCTTCAACCCAGCAGGAC 855
Db      770  GAAAGCTCTGGAGCTGCTGAAGAACGCGCATGCTAAGGCGGGCTACAC-----CGACAAG 823
Qy      856  ATCGCTCTTCTGCTGAGGCTTGTCTCTGAGTTCCTCAAGGACGCGCACCTACCACTTC 915
Db      824  ATTGTGATCGGCAATGACGCTGGCTGCTCGAGTTCTACAAAGGCGGCAAGTACGACCTG 883
Qy      916  GAA-----GGTGGCAGACCTCCGCAACCATCGGCGCAACGATGAGATGGCAACGTT 954
Db      884  GACTTCAAGTCTCCGACGACCCCGCGCTACATTTCTCCGAGCAGCTGGCGGACCTC 943
Qy      955  TAGCTGAGCTCGTTGACGGTACCAATCGTCTCCATCGAGACCCACTGCGAGGAGAT 1014
Db      944  TACAGGGGCTTCGTCAAGATTACCCAGTGTGTTCATCGAGACCCCTTTTGACCGAGAC 1003
Qy      1015  GACTGGGAGGTTTACACCAACCTCACGCAACCATCGGCGCAACGATGAGATGCTGTTGGC 1074
Db      1004  GACTGGGAGGCTGGACCAATTTCAAGCCAGACCAACG-----ATTCAAGTGGTGGGC 1057
Qy      1075  GACGACTTCTTCTGCTACCAACCTCGAGCGCTGAAAGAGGGGATCGCTAAGAGAGGCTGCC 1134
Db      1058  GACGACCTCAGGCTACCAACCCCAACGATCGCCAAAGGCTGTGGGCCAGAGTCTCTGC 1117
Qy      1135  AACTCCATCTGTTAAGGTGAACGAGTGGTACCTCGTACCTACCGAGACCTTCGAGCGCTGC 1194
Db      1118  AACTGCTGTGCTCAAAATCAACAGATCGGCTCGGTCCGATCCCGAGTCCCTGCGAGGATGC 1177

```

```

Qy 1195 GACATGGCTCACCGCGAGCTACACCTCCATGATGCCACCGTTCCTGGTGAGACCGAG 1254
Db 1178 AAGATGGCCAGAGCAATGCTGGGGCGTGTGCTAGCTCAGCATCGCTCCGGGAGACGGAG 1237
Qy 1255 GACACCACTTCTGACCTCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
Db 1238 GACACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
Qy 1315 CCAGCAGCTTCCGACCGCTGCGCAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTT 1374
Db 1298 CCGTGGCGCTCCGAGCGCTGGCCAGTACAAACAGCTGCTCCGATCGAGGAGAGCTG 1357
Qy 1375 GCGACCGCGGCTGCTACGAGCTGCGAGCGATTCACAGCTTTCAGGG 1424
Db 1358 GCGACCGCGGCTTCCGCGGCAAGAACTTCAGGCAACCCATCTGAGCG 1407

RESULT 5
CR679338
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CR679338 1623 bp mRNA linear HTC 19-AUG-2004
Tetraodon nigroviridis full-length cDNA.
CR679338
CR679338.1 GI:51177245
HTC; cDNA; full-length; Tetraodon nigroviridis.
Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 1623)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segrete@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1623
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Muscle"

ORIGIN

Query Match 24.3%; Score 383.6; DB 3; Length 1623;
Best Local Similarity 60.8%; Pred. No. 2.5e-92;
Matches 796; Conservative 0; Mismatches 454; Indels 60; Gaps 8;

Qy 160 ATCATGCACTATTCGCTCGGAAATTCGACCTCCGCGGTACCCAAACCGTCGAGGCA 219
Db 87 ATCTGAAGATCCACGCTCGGAAATATTCGACCTCCGCGGCAACCCCAACCGTCGAGGTC 146
Qy 220 GAGCTTTCTCGATACGCTTCCACGCTGTCGAGCTGTTCATCCGCGGATCCACC 279
Db 147 GACCTGTACCAAGAAGTCTGTTTACAGCTGCG---GTCCCCAGGCGGCTCCACC 203
Qy 280 GCGCTCCACGAGGCTCATGAGCTGCGTGACGGTGGCGA---TCGCTACCTGGGCAAGGC 336
Db 204 GGCATCTACGAGGCGCTGGAGCTCCGCGCAACGACAAACCCGCTACATGGGCAAGGT 263
Qy 337 GTTTTGAAGGCTTGAACGCTCAACGAAGATTCGGGACGAGCTCGCTGGCCTAGAG 396
Db 264 GTCTTAAAGCTGTGTGAGATATCAATAAATAATTCGACCTGCACTGGTGGCAAGGAC 323
Qy 397 GCTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTTGACCGCC 450
Db 324 GTCAACGCTTCTGACACGAGCAAAATTCGAAGCTGATCTGGAGATGGACGACCGCAC 383
Qy 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGCTGCTGCTGCTGCTGCT 510

```

```

Db 384 AACAAATCTAAGTTCGGGGCCAAACGCAATCCTCGGCGTGTCTCTGGTGTGTGCAAGGCT 443
Qy 511 GCTGCTGATTTCCGAGCGCTCCACACTGTTTCGGCTAGATCGGTGGACCAACAGC----- 563
Db 444 GGTGACGAGAGAAAGGCGTGGCCCTCTACCGCAATCCGCCACCTGGCGGCAACCC 503
Qy 564 --ACAGTTCCTTCAGTTTCAATGATGAACATCATCAACGGTGGCGCTCAAGCTGATCC 621
Db 504 CAAAGTCATCTCCCGTTCCTCGCTTCAACGTCATCAACGGCGCTCCACGAGGCAAC 563
Qy 622 GGTGTTGAGCTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGAGCTTCTCTGAGGCT 681
Db 564 AGCTGGCCATGACGAGTTTATGATCTCCGCGTGGAGCCAGCAGCTTCAAAGAGGCC 623
Qy 682 CTCCGCAACGGCGCGAGGTCTACACGCACTCAAGTCCGCTCATCAAGGAAAGAGGG- 738
Db 624 ATCGGCATCGGCGCGAAGTCTACCAACATCTGAAGACGTCATCAAGAGAGTACCGC 683
Qy 739 ---CTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCTTCCGTCGGCTCCACCGT 795
Db 684 AAGGACGCCCAACCAACGTTGGGAGACGAGGAGGCTTTGCCCCCAACATCTCTGGAGACAAG 743
Qy 796 GAGGCTCTTGACCTTATCGTTAAGCAATCGAAGGCTGGCTTCAACCGAGGCAAGGAC 855
Db 744 GAAGCTCTGGAGCTGCTGAAGAACGCAATCGTAAGCGCGGTACAC-----CGACAAG 797
Qy 856 ATCGCTCTTGTCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTC 915
Db 798 ATTGTGATCGCATGACGTTGGCTGCTCGAGTTCTCAAGGGCGCAAGTACGACTG 857
Qy 916 GAA-----GGTGGCGACGACTCCCGACGCTCCGACGCTGAGATGGCAAAAGCTT 954
Db 858 GACTTCAAAGTCTCCGACGACCCCGCGCGCTATCATCTCCCGAGCAGCTGGCGACCTC 917
Qy 955 TAGCTGAGCTCTGTTAGCGGTACCAATCGTCTCATCGAGACCCACTGCAGGAAGAT 1014
Db 918 TACAGGGGCTTCGTCAAAGATTACCCAGTGTGTCTCATCGAGACCCCTTTGACCAAGGAC 977
Qy 1015 GACTGGAGGGTTTACACCAACCTCACGCAACCATCGGCGCAAGGTTTCAGATCGTTGGC 1074
Db 978 GACTGGAGGGCTGGACCAATTTCAAGCAGCAGCAGCAGC-----ATTGAGTGTGGGC 1031
Qy 1075 GACGACTTCTTCTGTCACCAACCTGAGCGCTTGAAGAGGGCATTCGTAAAGAGGTGCC 1134
Db 1032 GAACGCTCTACGCTCACCAACCCCAACGATCGCAAGGCTGTGGCCCAAGAGTCTCTGC 1091
Qy 1135 AACTCATCTCTGTTAAGTGAACCGATCGGTACCTCACCGAGACTTCGACGCTGTC 1194
Db 1092 AACTGCTGTCTGCTCAAAGTCAAACGATCGGTTCGCTCACCGAGTCCCTGCAAGGATGC 1151
Qy 1195 GACATGGCTCACCGCGAGGCTACACCTCCATGATGTCACCGTTCGCGTGAACCGAG 1254
Db 1152 AGATGGCCAGAGAAAGGCTGGGGGTGATGTTGAGCCATCGTCCGGGAGACCGAG 1211
Qy 1255 GACACCACTTCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCT 1314
Db 1212 GACACCTTCTGCTGACCTGGTGTGCTGTGCAACCGGACAGATCAAGAGCGGCT 1271
Qy 1315 CAGGACGTTTCCAGCGTGTGCAAGTACAAACAGCTTCTCGCATCGAGCAGCTGCTT 1374
Db 1272 CCGTGGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCGGATTCGAGGAGAGCTG 1331
Qy 1375 GCGAGCGCGGCTCTACGAGGTTCGAGGCAATTCACCGCTTTCAGGG 1424
Db 1332 GCGACACGCGCGCTTCCGCGGCAAGAACTTACGGACCCCATCTGAGCG 1381

RESULT 6
BM321196
LOCUS
DEFINITION

```

```

BM321196 1505 bp mRNA linear EST 03-JAN-2002
rockefeller.0.899 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to enolase (EC 4.2.1.1), mRNA
sequence.

```



```

BM3211196
VERSION BM3211196.1 GI:18055602
KEYWORDS EST
SOURCE Mastigamoeba balamuthi
ORGANISM Mastigamoeba balamuthi
REFERENCE Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
AUTHORS 1 (bases 1 to 1505)
Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W.,
Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and
Philippe, H.
TITLE The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
PUBMED 11830664
COMMENT Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1505 Std Error: 0.00
POLYA=Yes.
FEATURES             Location/Qualifiers
     source            1..1505
                        /organism="Mastigamoeba balamuthi"
                        /mol_type="mRNA"
                        /strain="ATCC 30984"
                        /db_xref="taxon:108607"
                        /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                        /notes="syn: Phreatamoeba balamuthi"
ORIGIN
Query Match      24.28; Score 382.6; DB 4; Length 1505;
Best Local Similarity 59.98; Pred. No. 4.6e-92;
Matches 781; Conservative 0; Mismatches 469; Indels 53; Gaps 7;

564 TCACGAGGCCCTGGCATGGCCGCGAGAGTACCANAACTGGAAGCTGTNATCAAGA 623
731 AAAAGGGC-----CTGTCCACCGGACTTTGGCGATGA-GGGCGGCTTCGCTCCGTC 783
624 AGAGGTACGGCATGGACGCGACCAACGTGGCGACGAGGGGTTGCTTCGCCCCAACATC 683
784 GGCTCCACCGCTGAGGCTCTTGACCTTATGCTTAAAGCAATCAGAAAGGCTTGGCTTACC 843
684 CAGGCCAACACGAGGGCTTCGAGCTCATCGTTGAGGCCATCAAGCAGCGCGGTACACT 743
844 CCAGGCAAGACATCGCTCTGCTCTGGACGTTGCTTCCTCTGAGTTCTTCAAGGACGGC 903
744 GGCAAGATCGAGATCGGATGGACGTTGCCCTCTGTCGTTCTGGGACGCCAAGAGTCG 803
904 ACCTACACTTCGAAGGTGGCCAGACTCGCAGCTGAGATGGAACAACTTACGCTGAG 963
804 AAGTACGACCTCGGCTTCAAGGTCCCGCGGACGAAGAAGACCCCGACATGCTGNTCG 863
964 CTGTTGACGC-----GTACCAATCGTCTTCATC 993
864 GCGAGGNCCTTATCAAGCTCTACGAGGATGGACCTCGAAGTACCGGATCTGTCGATC 923
994 GAGGACCCACTGCGAGGAAGATGACTGGAGGTTTACACCAACCTCACCGCAACCATCGGC 1053
924 GAGGACCCCTTCGACGAGGAGACTGGGCCACTACACAGGTTACCGAGCTGATCCGC 983
1054 GACAAGTTTCAAGTCGTTGGCGACGACTTCTTCGTCAACCAACCTCAGCGCTTGAAGGAG 1113
984 AACAGGATCCAGATCGTCGGCGACGACCTGCTGTTCAACAAACCCCAAGCGCATCGTCGAG 1043
1114 GGATCGCTAAGAAAGGCTGCCAACTCATCTCTGTTAAGTGAACAGATCGGTACCTC 1173
1044 GCCCGCAACAAGAAAGGCTGCAACGCGCTGCTCAAGCTCAACAGATCGGCTCGGTC 1103
1174 ACCGAGACCTTCGACGCTGCGACATGGCTCACCGCGAGGCTACACCTCCATGATGTC 1233
1104 AGCGAGCGCTCGAGCGTCCG-CCTGGCCCGGAGGTCACTGGGGGCTCATGGTGTG 1162
1234 CACCGTTCGGTGAGACCGAGGACACCAATTCGTCGACCTGCGAGTTGCACCTCAACTGT 1293
1163 CACCGCTCGGGGAGACTGAGGACGCGTTCATCGCCGACCTCGTCTGCTGGGCTGC 1222
1294 GGCACATCAAGACTGGTCTCGACGCTTCGACCGTTCGCAAGTACACACGACTT 1353
1223 GGCAGATCAAGACTGGTCCCTCGGCTCGAGGCGCTCGCAAGTACCAACGAGTC 1282
1354 CTCGCGATCGAGCAGCTGCTTGGCGACGCGCGCTCTACGCGAG 1396
1283 CTCGCGATCGAGGAGGAGCTCGGGCCAAAGCCCACTACGCG 1325

RESULT 7
CR698859          1613 bp      mRNA      linear      HTC 19-AUG-2004
LOCUS            Tetraodon nigroviridis full-length cDNA.
DEFINITION       CR698859
ACCESSION        CR698859
VERSION          CR698859.1 GI:51196768
KEYWORDS         HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE           Tetraodon nigroviridis
ORGANISM         Tetraodon nigroviridis
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                  Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
                  Tetraodontidae; Tetraodontidae; Tetraodon.
TITLE            Genoscope.
JOURNAL          Direct Submission
COMMENT          Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
                  2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
                  [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
                  The sequences are based on single pass reads.
                  More information available at

```

```
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. .1613
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Eyes"

FEATURES
source
ORIGIN
Query Match 24.0%; Score 379; DB 3; Length 1613;
Best Local Similarity 61.0%; Pred. No. 4.4e-91;
Matches 800; Conservative 0; Mismatches 450; Indels 61; Gaps 9;

QY 160 ATCATGACGATTTTCGTCGCGAAATTTCTCGACTCCCGCGTAAACCCACCGTCGAGGCA 219
DB |||||
QY 81 ATCTGAAGATCCACGCTCGGGAATATTCGACTCCCGCGCAACCCACCGTGGAGTC 140
DB |||||
QY 220 GAGGTTTTCTGGATGAGGTTCCACGCGTGCAGGTGTTCCATCGGGGCATCCACC 279
DB |||||
QY 141 GACCTGTACACCAAGAGGTCTGTTTCAGAGCTGCG---GTCCCGAGCGCGCTCCACC 197
DB |||||
QY 280 GGGTCCACGAGGCTCATGAGCTCGCGTGCAGGCAAGCTTGCATCGCGGCAAGGCG 336
DB |||||
QY 198 GGCATCTACGAGCCCTGGAGCTCCGCGCAACGACAAACCCGCTACATGGGCAAGGT 257
DB |||||
QY 337 GTTTTGAAGGAGTTGAAACGTCACAGAAATTCGCGGACGAGCTCGCTGGCTTAGAG 396
DB |||||
QY 258 GTCTCTAAAGCTGTTGAGATATCAATAAAATATTCGACCTGCCTGCTGTTGGCAAGGAC 317
DB |||||
QY 397 GCTGACGATCAGCGC-----TCATCGACGAAGCAATGATCAAGCTTGATGGCACCGCC 450
DB |||||
QY 318 GTCAACGTTCTGGACACGAGGCAAAATTCGCAAGCTGATGCTGGAGATGGACGCGCACCGAC 377
DB |||||
QY 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGCTGCTGCTGCTGCTGCTGCT 510
DB |||||
QY 378 AACAAATCTAAGTTTGGGGCAACGCCATCTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
DB |||||
QY 511 GCTGCTGATTTCCGAGCGCTCCCACTGCTTCCGCTTACATCGGTGACCAACACGC----- 563
DB |||||
QY 438 GGTGACGACGAGAGGCGGTGCGCTCTACCGCACATCGCGACCTGCGCGGCAACCCC 497
DB |||||
QY 564 --ACAGTCTTTCAGTTCATGATGAACATCAACGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
DB |||||
QY 498 CAAGTATCTCTCCCGTTCGCGCTTTCGCGTATCAACGCGCGCTCCCAACGCGGCAAC 557
DB |||||
QY 622 GGTGTTGACGTTTCAAGGATTCATGATCGCTCCATCGTGCAGAGCTTCTCTGAGGCT 681
DB |||||
QY 558 AGCTGGCCATGACGAGGTTTCATGATCTGCGGTGGAGCCAGCAGCTTCAAGAGGCGC 617
DB |||||
QY 682 CTCGCGAACGCGCGGAGGCTTACACGCACTGAAGTCCGTCATCAAGGAAAGGCG--- 738
DB |||||
QY 618 ATGCGCATCGCGCGAGGCTTACCAACCTGAAGAACGTCATCAAGGAGAGTACGCG 677
DB |||||
QY 739 ---CTGTCCACGAGCTTGGGATGAGGGCGCTTCGCTCTCTCCGTCGCGCTCACCCGT 795
DB |||||
QY 678 AAGGACGCCACCAACGCTGGGAGACGAGGAGGCTTTGCCCGCAACATCTCTGGAGAACAG 737
DB |||||
QY 796 GAGGCTCTGACCTTATCGTTAAGGCAATCGAGAGGCTGCTTACCCCGGCAAGGAC 855
DB |||||
QY 738 GAAGCTCTGAGCTGCTGAAGACGCCATCGTAAGCGCGCTACAC-----CGACAG 791
DB |||||
QY 856 ATGCTCTTCTGTCGACGTTGCTTCTCTGAGTTCTTCAAGGACGCGCACTTACCACTTC 915
DB |||||
QY 792 ATTGTGATCGCATGAGAGTGGCTCTCGAGTTCTTCAAGGGCGCAAGTACGACCTG 851
DB |||||
QY 916 GAA-----GGTGGCAGACATCCGAGCTGAGATGGCAACGTT 954
DB |||||
QY 852 GACTTCAAGTCCCGGACGCGCGCTGCTATCTCTCCGAGAGCTGGCGGACCTC 911
DB |||||
QY 955 TAGGCTGAGCTGTTGAGCGGTACCCATCTGCTTCCATCGAGGACCCACTGCGAGGAGAT 1014
DB |||||
QY 912 TACAGGGGCTTCTGTAAGATTACCCAGTGTGCTTCATCGAGGACCCCTTTTGACGAGGAC 971
DB |||||

RESULT 8
CR650019
LOCUS 1615 bp mRNA linear HTC 18-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR650019
VERSION CR650019.1 GI:51146464
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1615)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequenage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1. .1615
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Query Match 24.0%; Score 379; DB 3; Length 1615;
Best Local Similarity 61.0%; Pred. No. 4.4e-91;
Matches 800; Conservative 0; Mismatches 450; Indels 61; Gaps 9;

QY 160 ATCATGACGATTTTCGTCGCGAAATTTCTCGACTCCCGCGTAAACCCACCGTCGAGGCA 219
DB |||||
QY 81 ATCTGAAGATCCACGCTCGGGAATATTCGACTCCCGCGCAACCCACCGTGGAGTC 140
DB |||||
QY 220 GAGGTTTTCTGGATGAGGTTCCACGCGTGCAGGTGTTCCATCGGGGCATCCACC 279
DB |||||
QY 141 GACCTGTACACCAAGAGGTCTGTTTCAGAGCTGCG---GTCCCGAGCGCGCTCCACC 197
DB |||||
QY 280 GGGTCCACGAGGCTCATGAGCTCGGTGACGTTGCGA---TCGCTACCTGGGCAAGGCG 336
DB |||||
QY 198 GGCATCTACGAGCCCTGGAGCTCCGCGCAACGACAAACCCGCTACATGGGCAAGGT 257
DB |||||
QY 337 GTTTTGAAGGAGTTGAAACGTCACAGAAATTCGCGGACGAGCTCGCTGGCTTAGAG 396
DB |||||
QY 258 GTCTCTAAAGCTGTTGAGATATCAATAAAATATTCGACCTGCCTGCTGTTGGCAAGGAC 317
DB |||||
QY 397 GCTGACGATCAGCGC-----TCATCGACGAAGCAATGATCAAGCTTGATGGCACCGCC 450
DB |||||
QY 318 GTCAACGTTCTGGACACGAGGCAAAATTCGCAAGCTGATGCTGGAGATGGACGCGCACCGAC 377
DB |||||
QY 451 AACAAAGTCCCGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
DB |||||
QY 378 AACAAATCTAAGTTTGGGGCAACGCCATCTGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
DB |||||
QY 511 GCTGCTGATTTCCGAGCGCTCCCACTGCTTCCGCTTACATCGGTGACCAACACGC----- 563
DB |||||
QY 438 GGTGACGACGAGAGGCGGTGCGCTCTACCGCACATCGCGACCTGCGCGGCAACCCC 497
DB |||||
QY 564 --ACAGTCTTTCAGTTCATGATGAACATCAACGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
DB |||||
QY 498 CAAGTATCTCTCCCGTTCGCGCTTTCGCGTATCAACGCGCGCTCCCAACGCGGCAAC 557
DB |||||
QY 622 GGTGTTGACGTTTCAAGGATTCATGATCGCTCCATCGTGCAGAGCTTCTCTGAGGCT 681
DB |||||
QY 558 AGCTGGCCATGACGAGGTTTCATGATCTGCGGTGGAGCCAGCAGCTTCAAGAGGCGC 617
DB |||||
QY 682 CTCGCGAACGCGCGGAGGCTTACACGCACTGAAGTCCGTCATCAAGGAAAGGCG--- 738
DB |||||
QY 618 ATGCGCATCGCGCGAGGCTTACCAACCTGAAGAACGTCATCAAGGAGAGTACGCG 677
DB |||||
QY 739 ---CTGTCCACGAGCTTGGGATGAGGGCGCTTCGCTCTCTCCGTCGCGCTCACCCGT 795
DB |||||
QY 678 AAGGACGCCACCAACGCTGGGAGACGAGGAGGCTTTGCCCGCAACATCTCTGGAGAACAG 737
DB |||||
QY 796 GAGGCTCTGACCTTATCGTTAAGGCAATCGAGAGGCTGCTTACCCCGGCAAGGAC 855
DB |||||
QY 738 GAAGCTCTGAGCTGCTGAAGACGCCATCGTAAGCGCGCTACAC-----CGACAG 791
DB |||||
QY 856 ATGCTCTTCTGTCGACGTTGCTTCTCTGAGTTCTTCAAGGACGCGCACTTACCACTTC 915
DB |||||
QY 792 ATTGTGATCGCATGAGAGTGGCTCTCGAGTTCTTCAAGGGCGCAAGTACGACCTG 851
DB |||||
QY 916 GAA-----GGTGGCAGACATCCGAGCTGAGATGGCAACGTT 954
DB |||||
QY 852 GACTTCAAGTCCCGGACGCGCGCTGCTATCTCTCCGAGAGCTGGCGGACCTC 911
DB |||||
QY 955 TAGGCTGAGCTGTTGAGCGGTACCCATCTGCTTCCATCGAGGACCCACTGCGAGGAGAT 1014
DB |||||
QY 912 TACAGGGGCTTCTGTAAGATTACCCAGTGTGCTTCATCGAGGACCCCTTTTGACGAGGAC 971
DB |||||
```

Db 198 GGCATCTACGAGGCCCTGGAGCTCGCGCAACACAAAACCCCGCTACATGGGCAAAAGGT 257
Qy 337 GTTTTGAAGGAGGTGAAACGTCACAGAAAGAAATCGCGCGAGAGCTCGCTGGGCTAGAG 396
Db 258 GTCTCTAAAGCTGTGAGAAATATCAATAAATAATATGCACCTGCACTGGTTGGCAAGGAC 317
Qy 397 GCTCAGCATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGATGGCACCGCC 450
Db 318 GTCACGCTTCGGACAGGCAAAATCGAAGCTGATGCTGGAGATGGACGGCACCGAC 377
Qy 451 AACAAAGTCCCGCTGGGTGCAACCGCAATCTTGGTGTTCATCGGCTGTGTGCAAAAGGCT 510
Db 378 AACAAATCTAAGTTCGGGGCAACGCCATCTGGGCGTGTCCCTGGCTGTGTGCAAGGCT 437
Qy 511 GCTGCTGATTCGGCAGGCTCCCACTGTTCCGCTTACATCGGTGACCAAAAGC-----563
Db 438 GGTGCAGCAGAGAAAGGGCGTCCACTCTACCGGCACATCGCGACCTGGCGGCAACCCC 497
Qy 564 --ACACGTTCTCCAGTTCCAATGATGAACATCATCAACGTTGGCGCTCAGCTGACTCC 621
Db 498 CAAGTCATCTCCCGCTTCGCTTTCAAGCTCATCAACGGCGGCTCCACAGCAGGAAC 557
Qy 622 GGTGTGACGTTACAGGAATTCATGCTCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCT 681
Db 558 AAGTGGCCATGCAAGGATTCATGATCTGCGGTGGGACGACGCTTCAAGAGGCC 617
Qy 682 CTCGCAACGGCGGAGGTTTACACGCACTGAAGTCCGTATCATCAAGGAAGAGGCG---738
Db 618 ATGGCATCGGCGCGAAGTCTACCAAACTGAAGAACGTCATCAAGAGGAAGTACGGC 677
Qy 739 ---CTGTCCACCGGACTTGGCGATGAGGGCGGCTTCTCTCTCGTGGCTCCACCGCT 795
Db 678 AAGGACCCCAACGCTGGGAGCAGGAGGCTTTCGCCCCCAACATCTGGAGAACAG 737
Qy 796 GAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGAC 855
Db 738 GAAGCTCTGAGTGTCTGAAGAACCGCATCGCTAAGGCGGGCTACAC-----CGACAAG 791
Qy 856 ATGCTCTCTCTGAGCGTGTCTCTCTGAGTCTTCAAGACGGCACCTACCACTTC 915
Db 792 ATTGTGATCGGATGAGCGTGGCTGCTCTGAGTCTTCAAGGGCGGCAAGTACGACCTG 851
Qy 916 GAA-----GGTGGCGAGCACTCCGCGAGTGGATGGCAAGCTT 954
Db 852 GACTTCAAGTCTCCGAGACCCCGCGCTACATCTCTCCGAGCAGCTGGCGGACCTC 911
Qy 955 TAGCTGAGCTCGTTGACGGTACCCAACTCGTCTCCATCGAGACCCACTGCAAGGAGAT 1014
Db 912 TACAGGGGCTTCGTCAAAGATTACCCAGTGGTGTCCATCGAGGACCCCTTTGACCAAGGAC 971
Qy 1015 GACTGGAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGTTTCAGATCGTTGGC 1074
Db 972 GACTGGAGGCTGGACCAATTTCAAGCAGCAGCAGC-----ATTGAGGTGGTGGG 1025
Qy 1075 GACGACTCTTCTGTCACCAA--CCCTGAGCGCTGAAGGAGGGGATCGCTAAGAGGCTGC 1133
Db 1026 GACGACCTCAGCGTCAACACCCCAACCGCATCGCAAGGGGTGTGSCCCAGAGTCTCTG 1085
Qy 1134 CAACTCCATCTCTGGTTAAGGTGAACAGATCGGTACCTCACCAGACACTTCAAGCCTGT 1193
Db 1086 CAACTGCTGCTGCTCAAAGTCAACAGATCGGCTCGGTCAACCGAGTCCCTGCGAGCATG 1145
Qy 1194 CGACATGGCTCACCGGCGAGGCTACACCTCCATGATGTCCACCGTTTCGGGTGAGACCGA 1253
Db 1146 CAAAGTGGCCAGAGCAACCGCTGGGCGGTGATGGTCAGGCATCGCTTCGGGAGACGGA 1205
Qy 1254 GGAACACCATTTGCTGACCTCGAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGC 1313
Db 1206 GGAACACTTTCATCGCTGACCTGGTGTGCTGTGCAACCGACAGATCAAGACGGGCGC 1265
Qy 1314 TCCAGACAGTTCCGACCGGTGTCGAAGTACACACGAGTCTTCCGATCGAGCAGTGTCT 1373
Db 1266 TCCGTCCCGCTCCGAGGGCTTGGCCAAAGTACAAACAGAGTCTCCGGATCGAGGAAGCT 1325

Qy 1374 TGGCGACGCGCGGCTCTAAGCAGGTGCGCAGCGCATTTCCACGCTTTTCAGGG 1424
Db 1326 GGGCGACCGAGCGCGTTCGCGCGCAAGACTTCAGGCACCCCATCTGAGGG 1376

RESULT 9
CR728153
LOCUS
DEFINITION
Tetraodon nigroviridis full-length cDNA.
ACCESSION
CR728153
VERSION
CR728153.1 GI:51226414
KEYWORDS
HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE
Tetraodon nigroviridis
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE
1 (bases 1 to 1637)
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
COMMENT
[E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
source
1..1637
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

ORIGIN
Query Match 24.0%; Score 378.6; DB 3; Length 1637;
Best Local Similarity 60.5%; Pred. No. 5.7e-91;
Matches 792; Conservative 0; Mismatches 459; Indels 58; Gaps 8;
Qy 160 ATCATGACGATTTCCGTCGCGAAATTCCTCGGCGTAACTCCCGGGTAAACCAACCGTCGAGGCA 219
Db 108 ATCTGAAATCCACGCTCGGGAATATTCGACTCCGCGGCAACCCACCGTGGAGGTC 167
Qy 220 GAGGTTTTCTGGATGACGGTTCACCGGTGCGAGGTGTTCCATCCGCGCATCCACC 279
Db 168 GACCTGTACCAAGAAAGGTCTGTACAGAGCTCGGTCC--CCAGCGCGCGCTCCACC 225
Qy 280 GCGTCCACGAGCTCATGAGCT--GCGTGACGGTGGCGATCGCTACCTGGGCAAGGGCG 337
Db 226 GGCATCTAGAGGCTTGGAGCTCCGCAACGCAAAACCCGCTACATGGGCAAGGTG 285
Qy 338 TTTTGAAGGCACTTGAAACCGTCAACGAAGAAATCGGCGACGAGCTCGCTGCGCTAGAGG 397
Db 286 TCTCTAAAGCTGTGAGAAATATCAATAAATAATTTGCACCTGCTGTTGGCAAGGAC 345
Qy 398 CTGACGATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGATGACACCGCA 451
Db 346 TCAAGCTTCTGGACGAGGCAAAATCGCAAGCTGATGCTGGAGATGGACGACCGCA 405
Qy 452 ACAAGTCCCGCTCGGTGCAAAACGCAATCTTTGGTGTTCATGCTGCTGTTCGCAAGGCTG 511
Db 406 ACAATCTAAGTACGGGGCAACGCTCTCTGGGCTGTCCCTGGCTGTGTGCTCAAGGCTG 465
Qy 512 CTGCTGATTCGCGAGGCTTCCACTGTTCCGCTATCATCGGTGGACCAACGC-----563
Db 466 GTGCAGCAGAGAGGGCGTGCCTCTACCGGCACATCGCGACCTGGCGGCAACCCCGC 525
Qy 564 -ACAGTCTTCCAGTTCCTCAATGATGAACATCATCAAGGTGGCGCTCAGCTGACTCG 622
Db 526 AAGTCATCTCCCGTTCCTCGCTTTCAAGCTCATCAAGCGCGGCTCCCAACGAGGCAACA 585
Qy 623 GTGTTGACGTTTCAGGAATTCATGATCGCTCCAAATCGGTGCGAGAGACCTTCTCTGAGGCTC 682

```
Db 586 AGCTGGCCATGCGAGGAGTTTCATGATCTGCGGTGGAGCCAGCAGCTTCAAGAGGCCA 645
Qy 683 TCAGCAACGCGCGGAGGTTCTACCAAGCAGTGAAGTCCGTCATCAAGGAAAGGCG--- 738
Db 646 TGGCATCGGCGCGGAGTCTACCAACACTGAAGACGTCATCAAGGAGAGTACGGCA 705
Qy 739 --CTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCGTCGGCTCCACCGGTG 796
Db 706 AGGACGCGCACCAAGTGGGAGACGAGGAGGCTTTGCCCCCAACATCTGGAGAACAGG 765
Qy 797 AGGCTTTGACCTTATCGTTAAGGCATCGAGAGGCTGGCTTCAACCCAGGCAAGACA 856
Db 766 AAGCTCTGGAGCTGCTGAAGAACGCCATCGCTAAGCGCGCTACAC-----CGACAAGA 819
Qy 857 TCCTCTTGTCTGCGAGCTTCTCTCTGAGTCTTCAAGGACGCGACCTACCACTTCG 916
Db 820 TTGTGATCGGCATGGAGCTGGCTCGCTCCGAGTTCTACAAGGGCGGAGTAGACCTGG 879
Qy 917 AA-----GGTGGCCAGCACTCCGAGCTGAGATGGCAAAAGTTT 955
Db 880 ACTTCAAGTCTCCCGACGACCCCGCGCTACATTTCTCCGAGCAGCTGGCGACCTCT 939
Qy 956 ACCTGAGCTCTGTGAGCGCTACCAATCGTCTCCATCGAGGACCCACTCAGGAAGATG 1015
Db 940 ACAGGGGCTTCGTCAAAGATTACCCAGTGTGTCCATCGAGGACCCCTTTGACCAAGACG 999
Qy 1016 ACTGGGAGGTTACACCAACTCAGCGCAACCATCGGCGACAGGTTTCAGATGTTGGCG 1075
Db 1000 ACTGGGAGGCTGGACCAATTTTACAGCAGCACAGC-----ATTGAGTGTGGCG 1053
Qy 1076 ACAGCTTCTTCTGTCACCAACCTGAGCGCTGAGGAGGAGCATCGCTTAAGAGGCTGCCA 1135
Db 1054 ACAGCTTACGGTCAACCAACCCCAAGCATCGCAAGGCTGTGGCCAGAGTCTCTGCA 1113
Qy 1136 ACTGCATCTGTTAAGGTGAACAGATCGGTACCTTCACCGAGACCTTCGACGCTGTG 1195
Db 1114 ACTGCTCTGCTGCTCAAAGTCAACAGATCGGCTCGGTCAACCGAGTCTCTCGAGCATGCA 1173
Qy 1196 ACATGGCTACCGCGAGGCTACACCTCATGATGTCCACCGTTCGGGTGAGACCGAGG 1255
Db 1174 AGATGGCCAGACAGCGCTGGGGGTGATGTCAGCCATCGCTCGGGGAGACGGAG 1233
Qy 1256 ACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGCTC 1315
Db 1234 ACACCTTCACTGCTGACCTGCTGCTGCTGTCGACCGGACAGATCAAGACGGCGCTC 1293
Qy 1316 CAGCAGTTCGACCGGTGTCGCAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTG 1375
Db 1294 CGTGGCGCTCCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCCGATCGAGGAGCTGG 1353
Qy 1376 GCACGCGCGCTTACGAGGTGCGAGCGCATTCACCGCTTCAGGG 1424
Db 1354 GCACACGCGCGCTTCCGCGGCAAGAACTTCAGGCAACCCCATCTGAGCG 1402
```

RESULT 10

```
CR685932
LOCUS 1551 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR685932
VERSION CR685932.1 GI:51183839
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage
```

```
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source 1..1551
        /organism="Tetraodon nigroviridis"
        /mol_type="mRNA"
        /db_xref="taxon:99883"
        /tissue_type="Eyes"
```

ORIGIN

```
Query Match 23.5%; Score 371.2; DB 3; Length 1551;
Best Local Similarity 60.4%; Pred. No. 5.7e-89;
Matches 783; Conservative 0; Mismatches 453; Indels 60; Gaps 8;

Qy 174 CGTCCGGAATTCCTGACTCCCGCGGTAAACCAACCGTCGAGCAGAGGTTTCTCGGA 233
Db 27 CGTCCGGAATTCCTGACTCCCGCGCAACCCACCGTGGAGTCAACCTGTACACAA 86
Qy 234 TGACGGTTCACCGGTGTCGAGTGTCCATCCGGCGCATCCACCGCGTCCACGAGGC 293
Db 87 GAAGGTCTGTTCAAGACTCG---GTCCCAGCGCGCTCCACCGGCTCTACGAGGC 143
Qy 294 TCATGAGCTGCGTG---ACGGTGGCGATCTGCTGGGCAAGGGCGTTTGAAGGCGT 350
Db 144 CTTGAGCTCCCGCAGCAACGATTAAACCGCTACATGGGCAAGGGGTCAAAGGGCGT 203
Qy 351 TGAACGCTCAACAGAAATTCGGGACGAGCTCGTGGCTTAGAGGCTGACGATACGCG 410
Db 204 TAAATATATCAATGAGTTTTTGGCCCCCGCTTGTGTAAACGAGACGCTCAACGTTCTGGA 263
Qy 411 CC-----TCATCGACGAGCAATGATCAAGCTTGTGGCAACCGCAACGCTCCCGCT 464
Db 264 CAGGCAAAATCGAAGCTGATGTGGAGATGGACGGCACCGCAACAAATCTAAGTT 323
Qy 465 GGGTCAAAACGCAATCTTGGTGTTCATGGTGTTCGAAAGGCTGCTGCTGATTCGCG 524
Db 324 CGGGCCACAGCCATCTTGGCGTGTCCCTGGCTGTGTGCAAGGCTGTCGACGAGAA 383
Qy 525 AGGCTCTCCACTGTTCCGCTACATCGGTGGACCAAAAG-----CACACGTTCTTCC 575
Db 384 GGGCGTGGCCCTCTACCGCCACATTCGCGACCTGGCGGCAACCCCAAGTCATCTCTCC 443
Qy 576 AGTTCAATGATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGACGTTC 635
Db 444 CGTTCCCGCTTTTCAACGTGATCAACGGCGCTCCACGAGGCAACAAAGCTGGCCATGCA 503
Qy 636 GGAATTCATGATCGCTCCAATCGGTGCAGAGACTTCTCTGAGGCTCTCCGCAACGCGC 695
Db 504 GAGTTTATGATCTCTGGGTGGAGCCAGCAGCTTCAAGAGGCGCATGCGATCGGCGC 563
Qy 696 GAGGTCTTACCAACGCTGAAGTCCGTCATCAAGGAAAGGGC-----CTGTCCACCGG 749
Db 564 CGAGGTCTACCAACACTGAAGAACGTCATCAAGGAGAGTACGGCAAGGACGCCACAA 623
Qy 750 ACTTGGCATGAGGGCGCTTCGCTCTTCCGTGGGTCCACCGGTGAGGCTCTTGACCT 809
Db 624 CGTGGAGAGCAGGAGGCTTTGCCCCCAACATCTCTGGAGAACAAAGGCTCTGAGCT 683
Qy 810 TATCGTTAAGGCAATCGAGAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTGCTCT 869
Db 684 GTTGAAGAACGCCATCGTAAGCGCGCTACAC-----CGACAAGATTGTGATCGGCAT 737
Qy 870 GGAAGTGTCTCTCTGAGTTCTTCAAGGACGCGACCTACCACTTCGAA----- 918
Db 738 GGAAGTGTCTCTCTGAGTTCTTCAAGGCGCGCAAGTACGACCTTGACTTCAAGTCCCC 797
Qy 919 -----GGTGGCCAGCATTCCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTCGT 968
Db 798 CGACGACCCCGCGCTTACATCTCTCCGAGCAGCTGGCGGACCTCTACAGGGGCTTCGT 857
```


QY 1313 CTCAGCAGCTTCGACCGGTGTCGAAAGTACAAACAGCTTCTCCGATCGAGCAGTGC 1372
 |||||
 Db 1293 CTCGTCGCGCTCCGAGCGGTTCGCAAGTACAAACAGCTGCTCCGATCAGGAGAGC 1352
 |||||
 QY 1373 TTGGCGAGCGCGGGGTCTACGAGGTGCGAGCGCATTTCCACGCTTTCAGGG 1424
 |||||
 Db 1353 TGGCGACACGAGCGCCCTTCGCGCGAAGAACTTCAGGCACCCCATCTGAGCG 1404
 |||||

RESULT 12
 CR656482
 LOCUS Tetraodon nigroviridis full-length cDNA. linear HTC 18-AUG-2004
 DEFINITION
 ACCESSION CR656482
 VERSION 1 GI:51152927
 KEYWORDS HTC: cDNA; full-length; Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.
 1 (bases 1 to 1582)
 Genoscope.
 Direct Submission
 Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 More information available at
 http://www.genoscope.cns.fr/tetraodon.
 FEATURES
 Location/Qualifiers
 1..1582
 /organism="Tetraodon nigroviridis"
 /mol_type="mRNA"
 /db_xref="taxon:99883"
 /tissue_type="Liver"

ORIGIN
 Query Match 23.2%; Score 366.2; DB 3; Length 1582;
 Best Local Similarity 60.4%; Pred. No. 1.3e-87;
 Matches 792; Conservative 0; Mismatches 458; Indels 61; Gaps 9;

QY 160 ATCATGCAGCTATTCGCTCCGAAATTCGACTCCCGGGTAACCCACCGTCGAGGCA 219
 |||||
 Db 67 ATCTGAAGATCCAGCTCGGGAATATTGATCTCCCGCGCAACCCACCGTGGAGTC 126
 |||||
 QY 220 GAGGTTTTCTGGATGACCGTTCCCAAGGTGTCGAGGTGTTCCATCCGCGCATCCACC 279
 |||||
 Db 127 GACCTGTACACCAAGAAAGTCTGTTTCAGAGTGG--GTCCCGAGCGCGCTCCACC 183
 |||||
 QY 280 GCGTCCAGAGCTCATGAGCTGCGTGGAGGGA---TGCTACCTGGGCAAGGC 336
 |||||
 Db 184 GGCATCTACAGGCGCTGGAGCTCCGCGCAACGACAAACCCGCTACATGGGCAAGGT 243
 |||||
 QY 337 GTTTTGAAGCAGTGAAGACGTCACAGAGAAATCGGACAGCTGCTGCGCTAGAG 396
 |||||
 Db 244 GTCTTAAAGCTGTGAGATATCAATAAATAATTCGACCTGCATGCTGGTTGGCAAGGAC 303
 |||||
 QY 397 GCTGAGCATCAGCGCC-----TCATCGACGAAGCAATGATCAAGCTTGTATGGCAGCC 450
 |||||
 Db 304 GTCAAGCTTCTGACCGAGGCAAAATCGCAAGCTGATGCTGGATGAGCGCACCGAC 363
 |||||
 QY 451 ACAAAGTCCCGCTGGGTGCAAGCAATCTTGGTGTTCATGCTGTTGCAAGGCT 510
 |||||
 Db 364 ACAAATCTAAAAACGGGGCAACGCCATCTCGGGGTGTCTCCCTGTGTGTGCAAGGCT 423
 |||||
 QY 511 GCTGCTGATTTCCGAGCGCTCCCACTGTTCCGCTACATCGGTGGAACAACGC----- 563
 |||||
 Db 424 GGTGAGCAGAGAGAGGCGTGGCACTCTACCGCACATCCCGACCTGGCGCGCAACCC 483
 |||||
 QY 564 --ACAGCTTCTCCAGTTCCAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCC 621
 |||||

Db 484 CAAGTCATCTCCCGCTTCCTCCGCTTTCAAGCTCATCAACGCGCGCTCCACGAGGAAC 543
 |||||
 QY 622 GGTGTTGAGCTTCAGGAATTCATGATCGCTCAATCGGTGCGAGAGCTTCTCTGAGGCT 681
 |||||
 Db 544 AGCTGGCCATGACGAGGTTTCATGATCTCTCGGTGGAGCCAGCAGCTTCAAAGAGGCC 603
 |||||
 QY 682 CTCGCAACGCGCGGAGGTCTACACGACCTGAAGTCCGCTCATCAAGGAAGAGGGC--- 738
 |||||
 Db 604 ATGGCATCGGCGCGGAGTCTACCAACCTGAAGACGTCATCAAGGAGAGTACGGC 663
 |||||
 QY 739 ---CTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCTTCCGTCGCTCCACCGT 795
 |||||
 Db 664 AAGGACGCCACCAACGCTGGGAGAGCGGTTCGCCCCCAACATCTCTGGAGAACAA 723
 |||||
 QY 796 GAGGCTCTTGACCTTATCTGTTAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGAC 855
 |||||
 Db 724 GAAGCCCTGGAGCTGCTGAAGAACGCTATCGTAAAGCGCGCTACAC-----CGACA 777
 |||||
 QY 856 ATCGCTTTGCTCTGGACGTTGCTCTCTGAGTCTTCAAGGACGCGCACTACCACTTC 915
 |||||
 Db 778 ATTGTCATCGGCATGAGCTGGCTGCTCTGAGTCTTACAAGGGCGGCAAGTACGACTG 837
 |||||
 QY 916 GAA-----GTGGCCAGCATCTCCGACGCTGAGATGGCAAAAGCTT 954
 |||||
 Db 838 GACTTCAAGTCTCCGACGACCCCGCGCTACATCTCTCCGAGCAGCTGGCGGACCTC 897
 |||||
 QY 955 TAGCTGAGCTCGTTACGCGTACCCCAATCGTCTCCATCGAGGACCCACTGCAGGAAGAT 1014
 |||||
 Db 898 TACAGGGGCTTCGTCAAAGATTACCCAGTGTGTCTCATCGAGGACCCCTTTGACAGGAC 957
 |||||
 QY 1015 GACTGGAGGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTTCAGATCGTTGGC 1074
 |||||
 Db 958 GACTGGAGGGGTGGACCAATTTTACAGCAGCAGCAGC-----ATTGAGTGGTGGC 1011
 |||||
 QY 1075 GACGACTTCTTCTGTCACCA--CCCTGAGCGCTTGAAGGAGGCGATCGCTAAGAGGCTGC 1133
 |||||
 Db 1012 GAACGCTCAGCGTCAACCAACCCCAAGCATCGCCAGGGTGTGSCCCAGAAAGTCTG 1071
 |||||
 QY 1134 CAACTCCATCTCTGTTAAGTGAACGATCGGTACCTCCCTCACGAGACCTTCCAGCTGT 1193
 |||||
 Db 1072 CAATCTCTGCTGCTCAAAAGTCAACAGATCGGCTCGGTACCGAGTCCCTTCGAGGATG 1131
 |||||
 QY 1194 CGACATGGCTCACCGCGCAGGCTACACCTCCATGATGTCCACCGCTTCGCGTGAGACCGA 1253
 |||||
 Db 1132 CAGATGGCCAGAGCAACGCTGGGGGTGATGTCAGGCATCGCTCCGGGAGAGCGA 1191
 |||||
 QY 1254 GGACACCACTTGTGACTCGACTGCACTCAACTGTGGCCAGATCAAGACTGGTGC 1313
 |||||
 Db 1192 GGACACCTTTCATGTTCTCTGTTGGTCTGTGTCACCGGACAGATCAAGACGGCGC 1251
 |||||
 QY 1314 TCCAGCAGCTTCGACCGGTGTCGCAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCT 1373
 |||||
 Db 1252 TCCGTGCGCTTCGAGCGCTTGGCCAAAGTACAAACAGCTGCTCCGATCGAGGAAGCT 1311
 |||||
 QY 1374 TGGCGAGCGCGGCTCTACGAGTGCAGCGCATTTCCACGCTTTCAGGG 1424
 |||||
 Db 1312 GGGCGACGAGCGCGTTCGCGCAAGAACTTCAGGACCCCATCTGAGCG 1362
 |||||

RESULT 13

BH770828

LOCUS

DEFINITION

LIMtag569 MG1363 Random Sequence Tag Library Lactococcus lactis

subsp. cremoris genomic, genomic survey sequence.

ACCESSION

BH770828

VERSION

BH770828.1

KEYWORDS

GSS.

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 2021)

BH770828 2021 bp DNA linear GSS 01-MAY-2002
 LIMtag569 MG1363 Random Sequence Tag Library Lactococcus lactis
 subsp. cremoris genomic, genomic survey sequence.

BH770828.1 GI:20373785

GSS.

Lactococcus lactis subsp. cremoris

Lactococcus lactis subsp. cremoris

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Lactococcus.

1 (bases 1 to 2021)

AUTHORS TITLE JOURNAL COMMENT

Bolotin, A., Ehrlich, S.D. and Sorokin, A.
Studies of genomes of dairy bacteria Lactococcus lactis
Sci. Aliments (2002) In press
Contact: Sorokin A
Genetique Microbienne
INRA

CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is enoB (95%)
Class: shotgun
High quality sequence start: 30
High quality sequence stop: 1993.

FEATURES source

Location/Qualifiers
1..2021
/organism="Lactococcus lactis subsp. cremoris"
/mol_type="genomic DNA"
/strain="MGI363"
/sub_species="cremoris"
/db_xref="taxon:1359"
/clone_lib="MGI363 Random Sequence Tag Library"
/note="Vector: pSGM2; Site 1: SmaI; Library of
chromosomal fragments of L.lactis strain MGI363 was
prepared by partial AluI digestion or by sonication."

ORIGIN

Query Match 21.7%; Score 341.8; DB 8; Length 2021;
Best Local Similarity 56.9%; Pred. No. 5.5e-81;
Matches 691; Conservative 0; Mismatches 512; Indels 12; Gaps 3;

159 AATCATGACGATTCCTCGTCGCGAAATCTCGACTCCCGCGGTAAACCCACCGTCGAGGC 218
|||||
709 AATCGAAACATTCACCGAAGAAATTTTGATTCAGAGGAATCCAAACCGTTGAGGT 768
|||||
219 AGAGTTTTCCTGATGACGGTTCCACAGGTGTCGAGGTGTTCCATCCGCGCATCCAC 278
|||||
769 AGATGTTAGATTGACAGATGGAATTTTGGACGCGGCGGAGTTCCTCAGGAGCATCAAC 828
|||||
279 CGGCTCCACGAGCTCATGAGTCGCTGAGCGGTGGCGATCGCTACCTGGCGAAGGCGT 338
|||||
829 AGAGATCGTGAAGCGGTGCAACTTAAGATGCGGAGCGGAGTTCGAAGGAAAGGCGT 888
|||||
339 TTGAGGCGAGTTGAAACGCTCAACGAAAGAAATCGCGACGAGCTCGCTGGCCTAGAGGC 398
|||||
889 TTCTAAGCTGTTGCAAAATGTTATGTTGAAATTTATGAGCCTTAAAGGCCAATCACC 948
|||||
399 TGAGGATCAGGCGCTCATCGAAGCAATGATCAAGCTTTGATGGCAACGCCAACAGTC 458
|||||
949 ATTCAATCAAGCAAAATTAGACCATCTGATGATTGAACCTTGATGGGACAAAAATAAGTC 1008
|||||
459 CCGCTGGGTGCAAAACGCAATCTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGA 518
|||||
1009 TCCTTAGTGGCAATGCAATATTTGGTGTTCATGGCTATTGCTGTCGCGCAGCTAA 1068
|||||
519 TTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGCCACACGTTCTCCAGT 578
|||||
1069 CAGTGAAGAAATCCCACTTTATGTTATCTTGTGGAGTAGATTGGAACCTTCTCAA-- 1126
|||||
579 TCCAATGATGAACATCATCAACGGTGGCGCTCAACGCTGACTCCGGTGTGACGTTCAAG 638
|||||
1127 -CCATTTTCAATGTCATCAATGGTGGAGTGCAACGCGAATTCAGGAATTCAGCTTCAAGA 1185
|||||
639 ATTTCATGATCGCTCCCAATCGGTGCAGAGACTTCTCTGAGGCTCTCGGCAACGCGCGGA 698
|||||
1186 ATTCTTGATTAACCCCTGTCAACGCTGCTAGTTTTCGGATGGTTTGAAGAAATCCGAAA 1245
|||||
699 GGTCTACCGCACTGAAGTCCGTCATCAAGGAAAAAGGCGCTGTCCACCGGACTTGGCGA 758
|||||
1246 TATCTATCATACTTGAAGAAAAATTTTAGCTGATAAAGGACTAGAAACGCGCGTAGGGGA 1305
|||||
759 TGAGGCGGCTTCGCTTCCTCCGTCGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAA 818
|||||

Db 1306 CGAGGGTGGTTTTGCTCCAAAACTTGGTTGCGACGGAATAATGCCATTGCTACATTTATCA 1365
Qy 819 GGCAATCGAGAGGCTGGCTTACACCCAGCAAGGACATCGCTCTTGCTCTCGACGTTGC 878
Db 1366 AGCAATTGAAAGAGCTGGTTATGTCGCCAGCGGAAGAAATGCGATTGCTATTGACCCGGC 1425
Qy 879 TTCTCTGAGTTCCTT-----CAAGACGGCACCTACCACTTCGAAGGTGGCAGCACTC 932
Db 1426 TTCTAGCGAATTTATGATGACAAAGAAAAAGTCTATCATTTTGAAGGTCAAAAGCTCAC 1485
Qy 933 CGCAGCTGAGATCGCAAAAGTTTACGCTGAGCTGTTGAGCGCTGACCCCAATCG--TCTC 989
Db 1486 TTCTAGCGAATTTGTTGACCTAATATGAGGGTTAGTTGAAAAATATCTCGCCCTAATCTC 1545
Qy 990 CATCGAGGACCCACTCGCAGGAAGATGACTGGGAGGTTTACCAACCACTCACCCGAACCAT 1049
Db 1546 TATTGAAGACGGATTTTCGGAAACATGATTGGGCTGGTTTTCGGGCTCAACAGAAAGTTCA 1605
Qy 1050 CGCGGACAAAGTTTCAGATCGTTGGCAGCAGCTTCTTCGTCACCAACCTGAGGCGCTGAA 1109
Db 1606 AGGTCAAAAAATTTCAATGTTGGTGTATGATATTTTGTCAAAATCCCGAAATCTTCAA 1665
Qy 1110 GGAAGGCATCGCTAAGAAGCTGCCAACTCCATCTCTGGTTAAGGTGAACCAAGATCGGTAC 1169
Db 1666 AGAAGAAATTCAAAAAGTGTAGCAATGCTATTTTGAATTAATAATCAAAATCCGGAC 1725
Qy 1170 CTTACCGAGACCTTCGACGCTGTGACATGGCTCAACCGCGAGGCTACACCTTCCATGAT 1229
Db 1726 AGTGACCGAAGCGATTTGAAGCGATTAGTTTGGCCAGAAAAAGCAGGTTATAAGACCATGAT 1785
Qy 1230 GTCCCAACCTTCGGTGGAGACCGGACACCACTTGTGCTGACCTCGAGTTGCACTCAA 1289
Db 1786 TTCTCATCTGTCAGGTGAAACCGTTGATAGTTTATTGCTGACTTTGCTATAGCCATGCA 1845
Qy 1290 CTGTGCGCAGATCAAGACTGGTGTCCAGCAGTTCGACCGGTGTCGCAAAAGTACAACA 1349
Db 1846 CGCTGACAAATAAAAAACAGGTTCAATGCGGAGAGTGAAGCGGTTGAAAAATATAATCA 1905
Qy 1350 GCTTCTCCGANTCGA 1364
Db 1906 ATTTTGGCTATTGA 1920

RESULT 14 CR733357

LOCUS CR733357 1407 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cdna.
ACCESSION CR733357
VERSION CR733357.1 GI:51231674
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1407)
Genoscope.

REFERENCE AUTHORS

Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information are available at
http://www.genoscope.cns.fr/tetraodon.

COMMENT

Location/Qualifiers

FEATURES source

1..1407
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="fish"

ORIGIN

```
Query Match      20.8%; Score 328.6; DB 3; Length 1407;
Best Local Similarity 59.8%; Pred. No. 1.9e-77;
Matches 687; Conservative 0; Mismatches 40; Indels 53; Gaps 6;

QY 317 ATCGCTACTCGGCAAGGCGTTTGAAGGCGAGTTGAAACGTCACGA-----AGAAAT 371
DB 13 ATGCTACTTGGGCAAGGGTCAAAGGGCGGTTAAATATGTCAATGAGTTTGGGCC 72
QY 372 CGGCGACGAGCTCGTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGAT 431
DB 73 CGCTTGTGTAAACAGGAGCTCAAGCTTCTGACAGCAAAATCGACAAGCTGATGCT 132
QY 432 CAAGCTTGATGGACCGCCAAACAAGTCCCGCTGGGTGCAAAACGCAATCTTGTGTTTC 491
DB 133 GGAGATGACGGCACCGCAACAATACTAAAGTTCGGGGCCAAACGCCATCTCTGGCGGTGTC 192
QY 492 CATGGCTTTGCAAGGCTGCTGATTCGCGAGGCTCCACTGTTCCGCTACATCGG 551
DB 193 CTTGGCTGTGCAAGGCTGTGTCAGCAGAGAGGGCGTGCCCTCTACCGCCACATCGC 252
QY 552 TGGACCAAAACG-----CACACGTTCTTCCAGTTTCCATGATGAACATCATCAACGG 602
DB 253 CGACTTGGCCGGCAACCCCAAGTCATCTCCCGTCCCGCTTCAACGTCATCAACGG 312
QY 603 TGGCGCTCACGCTGACTCCGCTGTGAGTTTCAAGAAATCATGATCGCTCCAAATCGGTGC 662
DB 313 CGGCTCCACGACGCAACAAAGCTGGCCATGAGAGTTTCAATGATCTCGCGGTGGAGC 372
QY 663 AGAGACTTCTCAGGCTCTCCCAACGGCGGAGGTCTACCAACGACTGAAGTCGT 722
DB 373 CAGCAGCTTCAAGAGGCGCATGCGCATCGGTGCGGAAGTCTTACCACAACTTGAAGAAGT 432
QY 723 CATCAAGGAAAGGCG-----CTGTCCACGGAATTTGGCGATGAGGCGGCTTCGTCTC 776
DB 433 CATCAAGGAGAGTACGCGAAGGAGCGCACCAAGTGGGAGACGAGGAGGCTTTGCCCC 492
QY 777 TTCGCTGGCTCCACCGTGAGGCTCTTGACTTATCGTTAAGGCAATCGAAGAGGCTGG 836
DB 493 CAACATCTGGAGAAACAGGAAGCTCTGGAGCTCTGAAGAAGCCATGCTAAGGCCGG 552
QY 837 CTTACCCCGGAGGAGACATCGCTCTGCTGAGAGTGTGCTTCTCTGAGTTCTTCAA 896
DB 553 CTACAC-----CGACAAGATTTGTGATCGGCGATGACGCTGGCTCGCTCCGAGTTCTCAA 606
QY 897 GGAGCGCACTTACCACTTCGA-----AGTGGCGCAGCACTCCG 935
DB 607 GGGCGGCAAGTACGACTGACTTCAAGTCTCCCGACGACCCCGCGCTACATCTCTCC 666
QY 936 AGCTGAGATGCAAAACGTTTACGCTGAGCTCGTTGACGCGTACCCAAATCGTCTCCATCGA 995
DB 667 CGAGCAGCTGGCCGACCTCTACAGGGGCTTCTCAAGAGATTACCCAGTGTGTTCCATCGA 726
QY 996 GGACCCACTGAGAAAGATGATCGGAGGGTTACCAACCTTACCGGAAACCATCGGCGGA 1055
DB 727 GGACCCCTTTGACCAAGACGATCGGAGGGGTGACCAATTTTACAGCCAGCACACG-- 784
QY 1056 CAAGGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACCTCGAGCGCTGAGAGGGG 1115
DB 785 -----ATTGAGTGTGGCGACGACCTTCAAGTCTCCCGACGACCCCAACCGCATCGCAAGGG 840
QY 1116 CATCGTAAAGAGGCTGCAACTCCATCTGTTTAAAGTGAACAGATCGGTACCTCTCAC 1175
DB 841 TGTGGCCAGAGTCTTGCACACTGCTGCTCTCAAGTCAACAGATCGGCTCGTTCAC 900
QY 1176 CGAGACTTTCGACGCTGTGACATGCTGCTACCGCGAGGCTTACACTTCCATGATGTCCCA 1235
DB 901 CGAGTCTCTGAGGATGCAAGATGGCCGAGCAACGCGCTGGGGCGTGTGTTGTCAGCCA 960
QY 1236 CGGTTCCGCTGAGACCGGAGACCAACATGCTGAGCTCGCAGTTGCACTCACTGTGG 1295
DB 961 TCGCTCCGGGAGACGAGGACACTTTCATCGCTGACCTTGTGCTGCTGTGTCACCGG 1020
QY 1296 CCAGATCAAGACTGGTGTCTCCAGCAGCTTCCGACCGGTGTGCGAAAGTACAACGAGCTTCT 1355
```

```
DB 1021 ACAGATCAATACGGGCGCTCCGTGCGCTCCGAGCGCTTGCCCAAGTACCAACGAGCTGCT 1080
QY 1356 CCGCATCGACGAGCTGTTTGGCGACGCGCGGCTTACGAGGTGCGACGCGATTTCCACG 1415
DB 1081 CCGATCGAGAGAGCTGGCGGACGAGGCGGCTTCCGCCGCAAGAACTTCAGGACACCC 1140
QY 1416 CTTTCAGGG 1424
DB 1141 ATCTGAGCG 1149

CR658829 1551 bp mRNA linear HTC 18-AUG-2004
Tetraodon nigroviridis full-length cDNA.
CR658829 GI:51155274
HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1551)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
(BE-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1..1551
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Muscle"

ORIGIN
Query Match 20.8%; Score 328; DB 3; Length 1551;
Best Local Similarity 57.9%; Pred. No. 2.8e-77;
Matches 740; Conservative 0; Mismatches 490; Indels 48; Gaps 7;

QY 160 ATATGACGCTATTCGCTCGGAAATTCGACTCCCGCGGTAAACCAACCGTCGAGGCA 219
DB 87 ATCATAAGATCCAGGCTCGGAGATCTCGACTCCAGAGGAAACCCACGGTGGAGTG 146
QY 220 GAGGTTTCTCTGATGACGTTCCCGCGGTGTCAGGTTTCATCGGCGCATCCACC 279
DB 147 GACCTGTGAGCGGCAAGGG---CCTGTTGAGAGCAGCGCTCCCGCGGGGCTCTACT 203
QY 280 GGCCTTCAAGAGGCTCATGAGCTGCGTGACGTTGGCGA---TCGCTACCTGGGCAAGGGC 336
DB 204 GGGGTCCATGAGCGGCTGGAGCTTCGTGATGGAGACAGAGTCTGCTACCTTGGCAAGGA 263
QY 337 GTTTTGAAGGAGTTGAAAACGTCACAGAAATCGGCGACGAGCTCGCTG-----GC 390
DB 264 ACCCTGAAGGCGGTGGACACGCTCAACAGGACATCGCCCCCAAGCTGATTGAGAAGAA 323
QY 391 CTAGAGGCTGACCATCAGCGCTCATCGAGAGCAATGATCAAGCTTGTATGGCACCACC 450
DB 324 TTCAGCGTTGTGAGGAGGAGAGATTGACCACTTTCATGCTGAGGCTGGATGGAACCGAG 383
QY 451 AACAAGTCCCGCTGGGTGCAACGCAATCTCTTGTGTTTCCATGCGTGTGCAAGGCT 510
DB 384 AACAAATCTCAGTTTGGCGCAACGCAATCTTGGGTGTCTCTGCGGCTGCTGCAAGGCC 443
QY 511 GCTGCTGATTTCCGAGGCTTCCCACTGTTCCGCTACATCGGTGGAGCAACACGACACG-- 568
DB 444 GGGCGGGGAGAGAGGCTTCTCTCTACCGCCACATCGCCGATCTCGCGGACACAA 503
```

```
Qy 569 -----TTCTTCAGTTTCAATGATGAACATCATCAACGGTGGCGCTACGCTGACTCC 621
Db 504 GACGTGATCTCCCGGTCTCTGCTTTCAACGTATCAATGGCGGAGCCACGCTGGGAAC 563
Qy 622 GGTGTTGACGTTTCAAGGAATTCATGATCGCTCCAATCGGTGCAGAGACCTTCTCTGAGGCT 681
Db 564 AAGCTGGCAATGACGAGTTTCAATGATTCCTCCCGTTGGAGGGCCAACTTCCACGAGGCC 623
Qy 682 CTCGCAACCGCGGAGGTTTACCAACGCTGAAGTCCGTCTCATCAAGGAAGAGGCG--- 738
Db 624 ATGAGGATCGGCGTGAAGTTTACCAACCTGAAGAGCGTGTCAAGGCAAGTACGGA 683
Qy 739 ---CTGTCCACCGGACTTGGCGATGAGGGGGCTTCCGTCTCTCCGTCCGCTCCACCCGT 795
Db 684 AAGAGCGCCACCAACGCTGGGAGAGCAAGGAGGCTTGCCCCCAACATCTCTGGAGAACAAC 743
Qy 796 GAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAAGGCTGGCTTCAACCCAGGCNAGGAC 855
Db 744 GAAGCTCTGAGCTGTGAAGTCAGCCATCGAAGGCGGCTACCCCGACAAGATCATC 803
Qy 856 ATCGCTCTTCTCTGAGCTGTGCTTCTCTGAGTTCTTCAAGGACGGCACTTACCACTTC 915
Db 804 ATCGGATGACGTGGCGGCTCCGAGTTCTATCGACGGGAAGTACGACCTGGACTTC 863
Qy 916 GA-----AGGTGGCCAGCACTCCGACGTGAGATGGCAACGTTTACGCT 960
Db 864 AAGTCCCCCATGACCCATCCAGGCACATCAGCGGGGAGAGCTGGGAGACCTGTACCGC 923
Qy 961 GAGTCTGTTGACGCTACCAATCGTCTCCATCGAGGACCCACTGCAGGAAGATGACTGG 1020
Db 924 AGCTTCATCCAGAATACCTGTTTCACTGTCATCGAAGACCCCTTTCACCAAGGATGACTGG 983
Qy 1021 GAGGTTTACCAACCTTCAACGCAACCATCGGCAACAAAGTTTCAAGTCTGATCGTTGGCGACGAC 1080
Db 984 GAGAACTGGGCCAAGTTTCAACGCTTCCACAGACATC-----CAGATCTGTGGAGATGAC 1037
Qy 1081 TTCTTCTGTCACCAACCTGAGCGCTGAGAGGGCATCGTAAAGAGCTGCGCACTCC 1140
Db 1038 CTGACGGTGACCAACCCCAAGAGATCCAGAGGGGTGCAAGAAAGGCTGTAACTGT 1097
Qy 1141 ATCTGTTTAAAGTGAACCAAGATCGGTACCTCACCGAGACCTTCCGACGCTGTCGACATG 1200
Db 1098 CTGCTGCTCAAGTCAACCAAGATCGGCTCAGTGCAGTCCATCAAGGCGTGTAACTG 1157
Qy 1201 GCTACCGCGCAGGCTACACCTCATGATGTCCACCGTTCCGGTGAAGCCGAGGACACC 1260
Db 1158 GCCCAGAGCAGTGGTTGGGGTGTGATGGTCAAGCCATCGCTCTGGAGAAACTGAGGACACC 1217
Qy 1261 ACCATTGCTGACCTCGCAGTTGGACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCA 1320
Db 1218 TTCACTCTCCACCTTGGTGGTGGACTGTGACCGGACAGATCAAGACGGGCGCTCCCTGC 1277
Qy 1321 CGTTCGACCGGTGTCGAAAGTACAAACAGCTTCTCCGCATCGAGCAGCTGCTTGGCGAC 1380
Db 1278 AGTTCGGAGGTTTGGCCAAAGTACAAACAGCTGATGAGATCGAGGAAGAGCTTGGGGAC 1337
Qy 1381 GCCGGGCTTACGAGGT 1398
Db 1338 AAGGCCAAGTTCGCCGT 1355
```

Search completed: September 28, 2005, 23:24:12
Job time : 4763.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 16:08:34 ; Search time 275.5 Seconds
(without alignments)
9372.213 Million cell updates/sec

Title: US-10-728-947-3

Perfect score: 1578

Sequence: 1 ggctgggagatggtagtt.....ctcaagcagggaacgtgctt 1578

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	4	US-09-860-768-3
2	1576.4	99.9	1578	4	US-09-860-768-1
3	632.6	40.1	4403765	3	US-09-103-840A-2
4	632.6	40.1	4411529	3	US-09-103-840A-1
5	550.4	34.9	1299	4	US-09-902-540-4780
6	550.4	34.9	27903	4	US-09-902-540-1235
7	532	33.7	1335	4	US-09-252-991A-4449
8	531.4	33.7	1347	4	US-09-252-991A-4158
9	479.6	30.4	1380	4	US-09-489-039A-4158
10	449.2	28.5	2363	4	US-09-818-780-22
11	440.8	27.9	3545	2	US-08-781-802-9
12	431.6	27.4	1305	4	US-09-583-110-1156
13	430.4	27.3	11384	3	US-08-961-527-45
14	428	27.1	1196	4	US-09-107-532A-156
15	418.6	26.5	1233	4	US-09-489-039A-2858
16	414	26.2	1511	3	US-08-847-065-18
17	414	26.2	1511	4	US-09-829-382-18
18	390.2	24.7	3513	3	US-09-058-260-1
19	390.2	24.7	4315	2	US-08-781-802-1
20	390.2	24.7	4315	3	US-08-694-078-1
21	389.6	24.7	1830121	4	US-09-557-884-1
22	389.6	24.7	1830121	4	US-09-643-990A-1
23	383.8	24.3	1371	4	US-09-543-681A-1918
24	381.6	24.2	1224	4	US-09-107-433-2383
25	374.8	23.8	1335	4	US-09-328-352-3143
26	367.2	23.3	1338	4	US-09-540-236-434
27	367.2	23.3	92407	4	US-09-596-002-36

28	359.6	22.8	1460	4	US-09-830-217-21	Sequence 21, Appl
29	359.6	22.8	1460	4	US-10-278-946-21	Sequence 21, Appl
30	351.2	22.3	1305	4	US-09-710-279-2221	Sequence 2221, Ap
31	351.2	22.3	1323	3	US-09-134-001C-2804	Sequence 2804, Ap
32	351.2	22.3	3830	4	US-09-710-279-3773	Sequence 3773, Ap
33	344.6	21.8	1539	4	US-08-956-171B-207	Sequence 207, App
34	344.6	21.8	1539	4	US-08-781-986A-207	Sequence 207, App
35	340.6	21.6	1230230	4	US-09-438-185A-1	Sequence 1, Appli
36	330.2	20.9	1230025	4	US-09-198-452A-1	Sequence 1, Appli
37	329	20.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
38	329	20.8	1664976	4	US-09-692-570-1	Sequence 1, Appli
39	308	19.5	1314	4	US-09-614-221A-509	Sequence 509, App
40	304.8	19.3	1314	4	US-09-614-221A-466	Sequence 466, App
41	304.8	19.3	1884	4	US-09-566-921-38	Sequence 38, Appl
42	301	19.1	934	3	US-08-961-083-183	Sequence 183, App
43	301	19.1	934	4	US-09-536-784-183	Sequence 183, App
44	299.6	19.0	2422	4	US-09-949-016-282	Sequence 282, App
45	299.6	19.0	2422	4	US-09-949-016-2488	Sequence 2488, Ap

ALIGNMENTS

RESULT 1

US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match	100.0%	Score	1578	DB	4	Length	1578
Best Local Similarity	100.0%	Pred. No.	0				
Matches	1578	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	GGCTGGGGATATGGGTAGTTTTCGCCACATAATTTCAACTGATTCCTCATCGAAACAAGA	60				
Db	1	GGCTGGGGATATGGGTAGTTTTCGCCACATAATTTCAACTGATTCCTCATCGAAACAAGA	60				
Qy	61	TTCTGTGCAACAATTTGGGTAGACGTGATTAAGACATTTTGATCAGTCAGTGAATATTTCTAG	120				
Db	61	TTCTGTGCAACAATTTGGGTAGACGTGATTAAGACATTTTGATCAGTCAGTGAATATTTCTAG	120				
Qy	121	TTAGTCTCCCAAGTTGGCATAGAGGCCACAGTGGTGAATATCATGACGATTCGTCGC	180				
Db	121	TTAGTCTCCCAAGTTGGCATAGAGGCCACAGTGGTGAATATCATGACGATTCGTCGC	180				
Qy	181	GAATTTCTGACTCCCGGTACCAACCGTCGAGCAGAGTTTTCCTGGATGACGGT	240				
Db	181	GAATTTCTGACTCCCGGTACCAACCGTCGAGCAGAGTTTTCCTGGATGACGGT	240				
Qy	241	TCCACAGGTGTCGAGGTTCCTCATCGGGCATCCACCGCGCTCCACGAGGTTCATGAG	300				
Db	241	TCCACAGGTGTCGAGGTTCCTCATCGGGCATCCACCGCGCTCCACGAGGTTCATGAG	300				
Qy	301	CTCGGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGAGTGAACAGTC	360				
Db	301	CTCGGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGAGTGAACAGTC	360				

```
QY 361 AACGAGAAATCGCGGACGAGCTCGCTGGGCTAGAGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAGAAATCGCGGACGAGCTCGCTGGGCTAGAGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTCGGTCGAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTCGGTCGAAACGCAATC 480
QY 481 CTTGGTGTTCATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
Db 481 CTTGGTGTTCATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
QY 541 CGCTACATCGGTGACCAAGACGACAGTTCCTCCAGTTCGAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGACCAAGACGACAGTTCCTCCAGTTCGAATGATGAACATCATCAAC 600
QY 601 GGTGGCGCTCACGCTGACTCCGGTGTGACGTTTACGAAATTCATGATCGCTCCAAATCGGT 660
Db 601 GGTGGCGCTCACGCTGACTCCGGTGTGACGTTTACGAAATTCATGATCGCTCCAAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGAGGTCTACCAACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACCGCGCGAGGTCTACCAACGCACTGAAGTCC 720
QY 721 GTCATCAAGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGGCGCTTCGCTCCTTCC 780
QY 781 GTGGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTAAGCAATCGAGAGGCTGGCTTC 840
Db 781 GTGGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTAAGCAATCGAGAGGCTGGCTTC 840
QY 841 ACCCGAGGCAAGACATCGCTCTTGCTCTGAGAGCTTCCTCTGAGTTCCTCAAGGAC 900
Db 841 ACCCGAGGCAAGACATCGCTCTTGCTCTGAGAGCTTCCTCTGAGTTCCTCAAGGAC 900
QY 901 GGCACCTTACCACTCGAGGTGGCGACACTCGCAGCTAGATGGCAACGTTAGCT 960
Db 901 GGCACCTTACCACTCGAGGTGGCGACACTCGCAGCTAGATGGCAACGTTAGCT 960
QY 961 GAGCTCGTTGACCGTACCAATCGTCTCCATCGAGGACCACTCGAGGAGATGACTGG 1020
Db 961 GAGCTCGTTGACCGTACCAATCGTCTCCATCGAGGACCACTCGAGGAGATGACTGG 1020
QY 1021 GAGGTTTACCAACTCTACCGCAACCATCGGCGACAAGGTTGATGCTGTTGGCGAC 1080
Db 1021 GAGGTTTACCAACTCTACCGCAACCATCGGCGACAAGGTTGATGCTGTTGGCGAC 1080
QY 1081 TTCTTGTGTCACCACTCGAGCGCTGAAGAGGCGATCGCTAAGAGGCTGCGCACTCC 1140
Db 1081 TTCTTGTGTCACCACTCGAGCGCTGAAGAGGCGATCGCTAAGAGGCTGCGCACTCC 1140
QY 1141 ATCTGTTAAGGTGAACAGATCGGTACCTCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db 1141 ATCTGTTAAGGTGAACAGATCGGTACCTCTCACCGAGACCTTCGACGCTGTCGACATG 1200
QY 1201 GCTACCGCGAGGCTACACTCGATGATGTCGCAACCGTTCGCGTGAACCGAGGAC 1260
Db 1201 GCTACCGCGAGGCTACACTCGATGATGTCGCAACCGTTCGCGTGAACCGAGGAC 1260
QY 1261 ACCATTGCTACCTCGCAGTTGACCTCAACTGTGGCCAGATCAAGTGGTGGCTCCAGCA 1320
Db 1261 ACCATTGCTACCTCGCAGTTGACCTCAACTGTGGCCAGATCAAGTGGTGGCTCCAGCA 1320
QY 1321 CGTTCGACCGTGTGCGAAAGTACAAACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCGACCGTGTGCGAAAGTACAAACAGCTTCTCCGCACTCGAGCAGCTGCTTGGCGAC 1380
QY 1381 GCCGGGCTTACGAGGTCCAGGGGATTCGACGCTTTTCAGGGCTTAATAAAGCGCTT 1440
Db 1381 GCCGGGCTTACGAGGTCCAGGGGATTCGACGCTTTTCAGGGCTTAATAAAGCGCTT 1440
```

```
QY 1441 TTTCGACGCCCGGTAACTCAAGGTTCCGGGCGCTCGTTGCTTACTACTTACTGTGT 1500
Db 1441 TTTCGACGCCCGGTAACTCAAGGTTCCGGGCGCTCGTTGCTTACTACTTACTGTGT 1500
QY 1501 GACTATGATCCGAGGATTTATGGCAAGCAGAAAGAACTATAAAGGCTTTGTTCCCTCT 1560
Db 1501 GACTATGATCCGAGGATTTATGGCAAGCAGAAAGAACTATAAAGGCTTTGTTCCCTCT 1560
QY 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 2
US-09-860-768-1
; Sequence 1, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-860-768-1

Query Match 99.9%; Score 1576.4; DB 4; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGGGATATCGGTAGTTTTCGCCACTAATTTCAACTGATTTGCCCTCATCGAAACAAGA 60
Db 1 GGCTGGGATATCGGTAGTTTTCGCCACTAATTTCAACTGATTTGCCCTCATCGAAACAAGA 60
QY 61 TTCTGCAACAATTTGGGTGTAGACGTGATTTGAAGACATTTGTATCAGTGAATTAATCTAG 120
Db 61 TTCTGCAACAATTTGGGTGTAGACGTGATTTGAAGACATTTGTATCAGTGAATTAATCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGACGTTATTCGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGACGTTATTCGCTCGC 180
QY 181 GAAATTTCTCGACTCCCGCGGTAAACCCAAACCGTCCGAGCAGAGGTTTTCTTGGATGACGGT 240
Db 181 GAAATTTCTCGACTCCCGCGGTAAACCCAAACCGTCCGAGCAGAGGTTTTCTTGGATGACGGT 240
QY 241 TCCACGGTGTCCAGAGTGTTCATCCGGCGATTCACCGCGCTCCACGAGGCTCATGAG 300
Db 241 TCCACGGTGTCCAGAGTGTTCATCCGGCGATTCACCGCGCTCCACGAGGCTCATGAG 300
QY 301 CTGCGTGAAGGTGGCGATCGTACCTTGGCAAGGGGTTTTGAAGCAGTTGAAACGTC 360
Db 301 CTGCGTGAAGGTGGCGATCGTACCTTGGCAAGGGGTTTTGAAGCAGTTGAAACGTC 360
QY 361 AACGAGAAATCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAGAAATCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTCGGTCGAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGAATGCGACCGCCAAAGTCCCGCTCGGTCGAAACGCAATC 480
QY 481 CTTGGTGTTCATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
```

[illegible]

Db 1561 CAAGCAGGNACTGTCTT 1578

RESULT 3
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match	40.1%;	Score 632.6;	DB 3;	Length 4403765;
Best Local Similarity	68.4%;	Pred. No. 8.4e-166;		
Matches	891;	Conservative	0;	Mismatches 409;
				Indels 3;
				Gaps 1;
Qy	140	AGGAGGCCACAGTGGCTGAAATCATGACGATATTCGCTCGCGAAATTCGACTCCGCG 199		
Db	1144576	AGGAGAACCCAGTGC CGATTTATTCGACGAGTTAGGGCCCGAGAGATCTCGATTCGCGCG 1144635		
Qy	200	GTAACCCACCGTCGAGGCAGAGTTTCTTGATGACCGTTCACCGGTGTCGCGAGGTG 259		
Db	1144636	GCAACCCGAGCGTGGAGGTGCGGCTTATTCACGGGACATTCGCGCCGGCCGCGG 1144695		
Qy	260	TTCCATCCGCGCGCATCCACCGCGCTCCACGAGGCTCATGAGCTCGCTGACCGTGGCGCATC 319		
Db	1144696	TGCGGTGCGGCGCTCGACCGGGGAGCAGCGCCGTGAGTTGCGGACGGCGCGCATC 1144755		
Qy	320	GCTACTGGCGCAGGGCGTTTGAAGCGAGTTGAAACGTCACGAGAAATCGCGCAGC 379		
Db	1144756	GCTACGCGCGCAAGGCGTGC AAAAAGCCGTGCAGGCTGTTCCTGTATGAGATCGGCCCGG 1144815		
Qy	380	AGCTCGCTGCCTAGAGGCTGACGATCAGCGCCTCATCGACGAGCAATGATCAAGCTTG 439		
Db	1144816	CCGTGATCGGACTCAACGCCGACGACAGCGATGGTCGACAGGCGCTGGTGAGACTAG 1144875		
Qy	440	ATGGCACCGCCAAACAGTCCCGCTGGGTGCAACGCAATCCTTGGTGTGTTTCATGGCTG 499		
Db	1144876	ACGGCACCCCGACAGTCCCGCTGGGGGCAACGCGATCTTGGGTGTCTCGCTCGCTG 1144935		
Qy	500	TTGCAAAAGGCTGCTGTGATTCGCGAGGCTCCCACTGTTTCGCTACATCGGTGGACCAA 559		
Db	1144936	TTGCAAGGCGGGCGGATTCGCGCGAGGTGCGGTTGTTCCGTTATGTGCGGGGGCCAA 1144995		
Qy	560	ACGCACAGTTCTTCAGTTTCCAAATGATGAACATCATCAACGTTGCGCTCAGCTTGACT 619		
Db	1144996	ACGCGCACATTCGCGCGGTACCGATGATGAACATCCTCAACGGCGGCGCACGCCGATA 1145055		
Qy	620	CCGGTGTGACGTTGACGGAATTCATGATCGCTCCAATCGGTGCAGAGACTTCTCTGAGG 679		
Db	1145056	CCGCTGTGCAATTCAGAGTTCACTGATGCGCAATTGGCGCGCCAGCTTCGTCGAGG 1145115		
Qy	680	CTCTCCGCACAGCGGCGGAGGTCTACACGCATGAAGTCCGCTCATCAAGGAAAGCGGC 739		
Db	1145116	CGTTGCGCTGGGCGCTGAGGTGTACACGCGCTCAAGTCGCTCTGAAAGAGGAGGGC 1145175		


```
Qy 740 TGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCGTCGGCTCCACCGTGGG 799
Db 1145176 TGTCCACCGGCTGGGCGACGAAGCGGCTTCGCCCGGATGTGCCGGACCAACCGCG 1145235
Qy 800 CTCTTGACCTTATCGTTAAGSCAAATCGAAGAGGCTGGCTTCAACCCAGGCAAGACATCG 859
Db 1145236 CGTTGGACCTGATCAGCGGGCCATCGAGTCGGCGGGCTTCGCGACCCGGCGCGAGCTGG 1145295
Qy 860 CTCTTGCTCTGGAGCTTGTCTCTCTGAGTCTTCAAGGAGGCGAC---CTACACATTCTG 916
Db 1145296 CGTGGGCCCTGGAGCGCGCGCCACCGAGTTCTTTCACCGACGCGACCGGCTACGCTCTCG 1145355
Qy 917 AAGTGGCCAGCACTCCGCGAGCTCAGATGGCAACGTTTACGCTGAGCTCGTTGACCGCT 976
Db 1145356 AGGCGACCAACCGTACCGCGAGACGAGATGACCGAGTTCTACCGGGCTGCTCGCGCCT 1145415
Qy 977 ACCCAATCGTCTCATCGAGGACCCACTGCGAGGAAGATGACTGGAGGGTTTACACCAACC 1036
Db 1145416 ACCCGCTGGTGCATCGAAGACCCACTGTCCGAAGACGATTGGGACGGCTGGGCGCGCG 1145475
Qy 1037 TCACCGCAACCATCGGCGACAGGTTTCAGATCGTTGGGAGGACTTCTTGTCTACCAACC 1096
Db 1145476 TGACGGGCTCGATCGGTGACCGGGTGCAATCGTTCGCGGACGACATCTTTGTCAACCAATC 1145535
Qy 1097 CTGAGCGCTGAAGGAGGCGATCGCTAAGAAGGCTGCCAACTCCATCTCGTTAAAGTGA 1156
Db 1145536 CCGAGCGCTCGAGGAGGCGATCGAACGGGCGTGGGCAATGCTGCTGCTCAAGTGA 1145595
Qy 1157 ACAGATCGGTACCTTCAACGAGACCTTCAACGCTGTCGATGCTTACCGCTACCGCGAGGCT 1216
Db 1145596 ACCAGATCGGAGCGTTGACCGGAGACACTCGACGCGGTTCACGCTGGCTTCAACGGCGAT 1145655
Qy 1217 ACACCTCCATGATGTCCTCCACCGTTCCGTCGAGACCGGAGGACACCACTTGTGACCTCG 1276
Db 1145656 ACCGACGATGATCAGTCAACGAGTGGCGAGACGAGGAGACACATGATCGCCGACCTCG 1145715
Qy 1277 CAGTTGCACTCAACTGTGGCGAGATCAAGACTGCTGTCAGACAGCTTCCGACCGTTCG 1336
Db 1145716 CGGTGGCCATCGGAGCGGCGAGATCAAGACGGGCGGCTGCTGCGAGTGAGGCGCTG 1145775
Qy 1337 CAAGTACACAGCTTCTCCGATCGAGCAGCTGCTTGTGGGACGCGCGGCTTACGCGAG 1396
Db 1145776 CAAATAACAACAGCTGCTCGGATCGAAGAGGCGCTTGGCGACGCGGCGGCTACGCGG 1145835
Qy 1397 GTGCGAGCGATTTCCACGCTTTTCAGGCGCTTAATAAAGCGCT 1439
Db 1145836 GCGACCTGGCATTTCTCGGTTTCGCTGCGAGAGCAATAGGT 1145878
```

RESULT 4

```
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
```

Query Match

40.1%; Score 632.6; DB 3; Length 441529;

```
Best Local Similarity 68.4%; Pred. No. 8.4e-166;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGGCTGAAAATCATGACGATATTCGCTCCGGAATTTCTCGACTCCCGCG 199
Db 1144551 AGGAGAACCCAGTGGCCGATTCAGACAGGTTAGGGCCCGAGAGATCTCTGATTTCCGCG 1144610
Qy 200 GTAAACCAACCGTCGAGGCGAGAGGTTTCTCGATGACGTTTCCACGGTGTCCGACGGTG 259
Db 1144611 GCAACCCGACGGTGGAGGTGAGGTGGCGCTTATTCGACGCGGACATTCGCCGCGGCGCG 1144670
Qy 260 TTCATCTCCGCGCATCCACCGCGTCCACGAGGCTCATGAGCTCATGAGCTGAGCGTGGCGATC 319
Db 1144671 TGCCTGTCGGCGGCTTCACCGGGGAGCACAGGCGCGTCGAGTTGCGGACGCGGCGGATC 1144730
Qy 320 GCTACCTGGGCAAGGGCGTTTGAAGGCGAGTTCAAAAACGTCAACGAAGAAATCGGCGACG 379
Db 1144731 GCTACGGCGGCAAGGGCGTGCAAAAAGCGTGCAGGCTGTTCTTGATGAGATCGGCCGG 1144790
Qy 380 AGCTCGTGGCCTTAGAGGCTGACGATCAGCCCTCATTCGACGAAGCAATGATCAAGCTTG 439
Db 1144791 CCGTCATCGGACTCAACGGCGGACGACAGCGATTGGTTCGACCAAGGCGCTGCTGACCTAG 1144850
Qy 440 ATGGCACCGCAACAAAGTCCCGCTCGGTGCAACGCAATCTCTGTTGTTTCCATGGCTG 499
Db 1144851 ACGGCACCCCGCAAGTCCCGCTCGGCGGCAACGCGATCTTTGGGTGCTCTCGCTCGCTG 1144910
Qy 500 TTGCAAGGCTGCTGCTGATTCGCGAGGCGCTCCACTGTTTCCGCTACATCGTGGACCAA 559
Db 1144911 TTGCCAAGGCGGCGCGGATTCGCGGAGCTGCGGTTGTTCCGTTATGTGCGGGGCCAA 1144970
Qy 560 ACGCACAGCTTCTTCAGTTCCAATGATGAAATCATCAACGTTGGGCTCAGCTGACT 619
Db 1144971 ACGCGACATTTCTGCCGCTAAGATGAACTCCTCAACGCGCGCGACACGCGGATA 1145030
Qy 620 CCGCTGTTGACGTTAGGAAATTCATGATCGTCCAATCGTGGCGCAATTTGGCGGCCAGCTTCTTGAGG 679
Db 1145031 CCGCTGTCGACATTCAGAGGTTCACTGTTGGCGCAATTTGGCGGCCAGCTTCTTGAGG 1145090
Qy 680 CTCTCCGAACGGCGCGAGGCTTACCAACGCTGAAAGTCCGTTCAATCAAGAAAGGGCC 739
Db 1145091 CGTTGGCTGGGCGCTGAGGTTGACCAACGCTCAAGTCCGTTCTGAAAAGAGGAGGGC 1145150
Qy 740 TGTCCACCGGACTTGGCGGCTTGGCTCTTCCGTCGCTCCAGCTCCACCGTGAAG 799
Db 1145151 TGTCCACCGGCTTGGCGGACGAAGGCGCTTGGCCCCGATGTGGCGCGGACCAACCGCG 1145210
Qy 800 CTCTTGACCTTATCGTTAAGGCAATCGAAGAGGCTGGCTTCAACCCGAGGCAAGGACATCG 859
Db 1145211 CGTTGACCTGATCAGCCGGGCGCATCGAGTCCGCGGCTTGGACCGGCGCGAGCTGG 1145270
Qy 860 CTCTTGCTCTGAGCGTTGTTCTCTGAGTTCTTCAAGACGCGCAC---CTACCACTTCG 916
Db 1145271 CGCTGGCCCTGGACGCGGCGGCGACCGAGTTCTTCAACGACGCGCACCGGCTAGCTCTTCG 1145330
Qy 917 AAGTGGCGAGCACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACGCT 976
Db 1145331 AGGCGACCAACCGTACCGGAGACCATGACCGAGTTCTACGCGGCGCTGCTCGGCGCT 1145390
Qy 977 ACCCAATCGTCTCCATCGAGGACCCCACTGACGAAGATGACTGGGAGGGTTACACCAACC 1036
Db 1145391 ACCCGCTGTTGCTGATCGAAGACCCCACTGTCGGAAGACGATTGGGAGCGCTGGGCGCGCG 1145450
Qy 1037 TCACCGCAACCATCGGCGACAAAGTTTCAGATCGTTGGCGACGACTTCTTCTGTCACCAACC 1096
Db 1145451 TGACGGCTCGATCGGTGACCGGGTGCAATCGTTCGCGGACGACATCTTTGTCAACCAATC 1145510
Qy 1097 CTGAGCGCTGAAGGAGGCGCATCGCTAAGAAGCGTCCCAACTCCATCTGTTTAAAGTGA 1156
Db 1145511 CCGAGCGGCTCGAGGAGGCGCATCGAACGGGCGTGGCAATGCGTTGCTTCAAGTGA 1145570
Qy 1157 ACAGATCGGTACCTTCAACCGGAGACTTTCGACGCTGTCGATGCTCAACGCGAGGCT 1216
```

Db 1145571 ACAGATCGGAGCTTTCAGCCGAGACATCGACCGGGTCAACGCTGGCTCAACACGCGCGAT 1145630
 Qy 1217 ACACCTCCATGATGTCACCGTTCGCGTGGAGAGGAGACACACCATTTGCTGACCTCG 1276
 Db 1145631 ACCGACGATGATCAGTCAACGAGTGGGAGAGGAGGACACCATGATCGCCGACCTCG 1145690
 Qy 1277 CAGTTGACCTCAACTGTTGGCCAGATCAAGAGTGGTGTCCAGACAGTTTCGACCGTTCG 1336
 Db 1145691 CGGTGGCCATCGGACGCGGCGAGATCAAGACGGGCGGCTTCTCGCAGTGAGCGCTCG 1145750
 Qy 1337 CAAAGTACACCGCTTCTCCGATCGAGAGCTGCTTGGGAGCGCGCTTACCGAG 1396
 Db 1145751 CAAATACACCGCTTCTCCGATCGAGAGGCTTGGGAGCGCGCTTACCGAG 1145810
 Qy 1397 GTGCGACGCAATTCACCGCTTTCAGGCTTAAATAAGCGCT 1439
 Db 1145811 GCGACCTGGCAATTCCTCGTTCGCTGCGAGAGCAATAGGT 1145853

RESULT 5
 US-09-902-540-4780
 ; Sequence 4780, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR FILING DATE: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 4780
 ; LENGTH: 1299
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-4780

Query Match 34.9%; Score 550.4; DB 4; Length 1299;
 Best Local Similarity 65.6%; Pred. No. 1.6e-144;
 Matches 841; Conservative 0; Mismatches 426; Indels 15; Gaps 2;

Qy 152 TGCTGAATCATGCAGTATTCGCTCGCGAAATTCGACTCCCGGTTAACCCACCG 211
 Db 2 TGACCGAGATTTCTCAGATTCGCGCGGAGTGTGACTCCCGGTGGCAACCCGACCG 61
 Qy 212 TCAGAGGAGAGGTTTCTCGATGACGGTTCACCGGTTCGAGGTGTTCCATCCCGCG 271
 Db 62 TGGAGGGGAGGTTCCAGCTTTCGCGGGGGCGCTCGTGGCGGGCGCGGTCCGCTCGGTG 121
 Qy 272 CATCCACCGGGTCCACGAGCTCATGAGTGTGCTGACGTTGGC---GATCGCTACCTGG 328
 Db 122 CGTCCACCGGTGAGCATGAGCCATCGAGTTCGCGACGCGGCAACGATCGCTACCTGG 181
 Qy 329 GCAAGGGCGTTTGAAGCGAGTTGAAACGTTCAACGAGAAATCGGCGAGCTCGCTG 388
 Db 182 GCAAGGGCGTGCAGAGGCGGTGAAGAACGTGTTGTCGAGCTGCTCGCGCCCGCGTGTGG 241
 Qy 389 GCCTTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGCTTGTATGGCACCG 448
 Db 242 GCATGGACGGCGGATCAGTTCGCGGTGACCGAGCGGATGCTGGAGCTGGACGCGACCG 301
 Qy 449 CCAACAGTCCCGCTCGGTTGCAACCAATCTTGGTGTTCATCGCTGTTGCAAGG 508
 Db 302 CCACCAAGGGCAAGCTGGGCGCAACCCATCTTCGCGGTGTCCATGGCGCGCTCGCG 361
 Qy 509 CTGCTGCTGATTCGCGAGGCTTCCATGTTCCGCTACATCGGTGGACCAACGCCACG 568
 Db 362 CCGCGCGGAGCGGACGCGGCTGCTGTAACCGTACGCTGCGGGCGGTGAGCGCGGCA 421

Qy 569 TTCTTCAGTTCCAAATGATGAACATCATCAACGCTGCGCTCACGCTGACTCCCGTGTG 628
 Db 422 CCCTGCGGTGCGCTGATGAACATCTCAACGCGCGCGCACCGGACACCCGCGTGG 481
 Qy 629 ACCTTCAGGAATTCATGATCGCTCCATCGGTGCGAGAGACTTCTCTGAGGCTTCGCGCA 688
 Db 482 ACGTGACGAGGATTCATGCTGCTGCGCGCGCTTCTCTGCGGAGGCGCTGCGCT 541
 Qy 689 ACAGCGCGGAGGTTCTACACGCACTGAAGTCCGTCATCAAGGAAAGAGGCGCTGCTCCACCG 748
 Db 542 GGGGCGCGGAGGTTTCCACGCGCTGAAGAATTTCTCAAGGGCGGCAAGCTGCGCACGG 601
 Qy 749 GACTTGGCGATGAGGGCGGCTTCCGCTCTTCGCTGCGCTCAACCGGTGAGGCTTTCGACC 808
 Db 602 GCGTGGCGAGAGGGCGGCTACGCCCGGACCTGCGCGGCAACGAAGAGCGCTGAAGC 661
 Qy 809 TTATCGTTAAGGCAATCGAAGAGGCTGCGCTTCAACCGAGGCAAGGACATCGCTCTTGTCTC 868
 Db 662 TCATCATTGGAGGCGCATCGACCGCGGCTTCAAGCGCGGTGAGCAGCTTCTCTGCGGCC 721
 Qy 869 TGGACGTTGCTTCTCTGAGTTCCTTCAAGGACGCGCACTTACCACTT-----CG 916
 Db 722 TGGACGTTGCGGCGGCGAGGAGTTCCTTCAAGGCGGCGAGAGTACAGCTGAAGGGCG 781
 Qy 917 AAGTGGCCAGCACTCCGAGCTGAGATGGAACGTTTACGCTGAGCTCGTTGACGCT 976
 Db 782 AGGCAAGGAGTACGACTCGACCGGCTGCTGAGTACTACCGGGGCTTCTCCGAGCGCT 841
 Qy 977 ACCCAATCGTTCATCGAGGACCCACTGCAAGGAGTGTGAGGAGGTTTACACCAACC 1036
 Db 842 ACCCAATCATCTCATCGAAGCGGCTGCGGAGGATGATGCGGAGGCTTGGGAGGCT 901
 Qy 1037 TCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGGCGAGCACTTCTTCGTCAACCAACC 1096
 Db 902 TCACCGAGCGCTGGGTTCCGCGATGCGAGTGGTGGTGAGCACTTCTTCGTCAACCAACC 961
 Qy 1097 CTGAGCGCTTGAAGGAGGCGATCGCTAAGAGGCTGCCAATCCATCTCGTTAAGGTGA 1156
 Db 962 TGGAGCGCTTGGCGCGCGGCGATCGAGACGGGCAACGCTTCTTCGTGAAGGTGA 1021
 Qy 1157 ACAGATCGGTACCTTCCGAGACCTTTCGACGCTGCGAGATGCTTCCGCGCGAGGCT 1216
 Db 1022 ACCAGATTTGCAACCTTTCGAGGAGCGTTCGACGCGCTGCGCATGCGCGCGCTT 1081
 Qy 1217 ACACCTCCATGATGCTCCACCGCTTTCGCTGAGACCGGAGGACACCACTTCTGACCTCG 1276
 Db 1082 ACAGTCCGTTCATGAGCACCGCTTCCGCGAGACCGGAGGACACCACTTCCGCGCGCTGG 1141
 Qy 1277 CAGTTGCACTCAACTGTTGGCCAGATCAAGACTGCTGCTCCAGCAGCTTTCGAGCCTGTCG 1336
 Db 1142 CCGTGGCGCTGGACTGCGGCGAGATCAAGACGGGTTCCGCGTCCCGCTCCGACCGCGTGG 1201
 Qy 1337 CAAAGTACACAGCTTTCGCGATTCGAGCAGCTGCTTGGGAGCGCGGCTTACCGAG 1396
 Db 1202 CCAAGTACAAACGATTCGCTGCGATCGAAGCGAGCTTGGGGCGCGCGCTACGCTG 1261
 Qy 1397 GTGCGAGCGCATTCACCGCTT 1418
 Db 1262 GCAAGTGGTTCGCGCGCTT 1283

RESULT 6
 US-09-902-540-1235/c
 ; Sequence 1235, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1235
; LENGTH: 27903
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-502-540-1235

Query Match 34.9%; Score 550.4; DB 4; Length 27903;
Best Local Similarity 65.6%; Pred. No. 8e-144;
Matches 841; Conservative 0; Mismatches 426; Indels 15; Gaps 2;
QY 152 TGGCTGAATCATGCAGATTTTCCTCGATGACGGTTCCACGGTGTCCAGGTGTTCCATCCGGCG 211
Db 19948 TGACCCAGAGATTTCAGATTCTGGCGCGCAAGTGTCTCGACTCCCGTGGCAACCCGACCG 19889
QY 212 TCAGGCGAGAGGTTTTCCTCGATGACGGTTCCACGGTGTCCAGGTGTTCCATCCGGCG 271
Db 19888 TGAGGCGGAGGTTCCAGCTTTCGCGGGGGCGCTCGTGGCCGCGCGCGGTGCCGTCCGGTG 19829
QY 272 CATCCACCGCGCTCCACGAGGCTCATGAGCTGGGTGACGGTGGC--GATCGCTACCTGG 328
Db 19828 CGTCCACCGGTGAGCATGAGGCCATCGAGCTGCGGAGCGGACAAAGCATCGCTACCTGG 19769
QY 329 GCAAGGGCGTTTGAAGGCAAGTTGMAAACGTCAACGAAGAAATCGGGCGAGAGTCCGCTG 388
Db 19768 GCAAGGGCGTGCAGAGGCGGTGAAGAACGTGTGTGACGTGCTCGCGCCCGCGCTGGTGG 19709
QY 389 GCCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGGCACCG 448
Db 19708 GCATGACGCGGGATTCAGTTCCGGTGGACGACGAGGATGCTGGAGCTGGAGCGCACCG 19649
QY 449 CCAACAGTCCCGCTGGGTGCAAAACGAATCCTTGGTGTTCATGGCTGTTGCAAGG 508
Db 19648 CCACCAAGGCAAGCTGGCGCGCAACGCCATCTCGGGGTGTCATGGCGCGCTCGCG 19589
QY 509 CTGCTGCTGATTCGGCAGGCTCCCACTGTTCCGTATACATCGGTGGACCAAAACGACG 568
Db 19588 CCGCGCGGACGGCGACGGGCTGCGCTGTACCGTACGTGGGCGGGGTGACGGCGGCA 19529
QY 569 TTCTTCAGTTCCAATGATCAATCATCAACGCTGCGCTCAGCTGACGTCCGGTGTG 628
Db 19528 CCTGCGGTCGCTGATGAACATCTCAACGGCGCGGACGCGGACACCGCGGTG 19469
QY 629 ACCTTCAGGAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCA 688
Db 19468 ACCTGACGAGTTTCATGGTGTGCGCGCGGTGCTCTCTCTTCGCGAGGGCTCGGCT 19409
QY 689 ACGCGCGGAGGTTTACCAACGACTGAAGTCCGTTCATCAAGGAAGAGGGCTGTCCACCG 748
Db 19408 GGGCGCGCGAGGTGTTCCACGCGCTGAAGAAGATTCTCAAGGGCGCGCAAGCTGGCCACGG 19349
QY 749 GACTTGGCGATGAGGGCGGCTTCGCTCTCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACC 808
Db 19348 GCGTGGCGACGAGGGCGGTGACCCCGGACCTGCGCGGACGAAGAGGCGCTGAAGC 19289
QY 809 TTATCGTTAAGGCAATCGAAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGCTC 868
Db 19288 TCATCATGAGGCGCATCGACGAGCGGGCTTCAAGGCGCGGTGAGCAGCTCTTCTGGGCC 19229
QY 869 TGGACGTTGCTTCTCTGAGTCTTCAAGGACGCGACCTACCACTT-----CG 916
Db 19228 TGGACGTTGGCGGCGAGGAGTTCTTCGACAAAGGCGACGAAGAATCAAGCTGAAGGGCG 19169
QY 917 AAGGTGGCCAGCACTCCGACGCTGAGATGGCAAAAGTTTACGCTGAGCTGTGTTGACCGGT 976
Db 19168 AGGGCAAGGATACGACTCGACCGGCTGCTCGAGTACTACGGGGCTCTCCGAGCGCT 19109
QY 977 ACCCAATGCTTCATCGAGGACCCCACTGCGAAGAGATGCTGGAGGGTTTACCAACAC 1036
Db 19108 ACCCCATCATCTCCATCGAAGCGGATGCGGAGGATGATGCTGGAGGGCTGGAAGAAGC 19049

RESULT 7

US-09-252-991A-4449

; Sequence 4449, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4449

; LENGTH: 1335

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4449

Query Match 33.7%; Score 532; DB 4; Length 1335;
Best Local Similarity 64.8%; Pred. No. 2.5e-139;
Matches 847; Conservative 0; Mismatches 440; Indels 21; Gaps 3;
QY 131 AGTTGGCATAGGAGGCGACAGTGGCTGGAATCATGACGCTATTCGCTCGGCAAAATTCG 190
Db 26 ACTTCGGAGTGTTAACAAGAAATGGCAAGATCGTCGACATCAAGGGCGGTGAGTCCCTGG 85
QY 191 ACTCCCGCGGTAAACCAACCGTCGAGCGAGAGGTTTTCTTGGATGACGTTTCCACCGGTG 250
Db 86 ACTTCGGCGCAACCTTACCGTTGAAGCGGACGTGATCTTGGACACGGCATCGTCGGCA 145
QY 251 TCGCAGGTGTTTCATCCGCGCATCCACCGCGCTCCACGAGGTTCAGAGGCTCATAGCTCGTGACG 310
Db 146 GCGCTTCGCGGCTTCCTCGGTGCTTCACCGGTTCCCGCGAGGCCCTCGAGCTGCGCATG 205
QY 311 GTGGCA---TCGCTACCTGGGCAAGGGCGTTTTGAAGGAGTTGAAAAAGTTCACGAAG 367
Db 206 GCACAAGAGCCCTTACCTGGGCAAGGGCGTGTCTGAAGCCGTGGCCCAACATCAACCGCC 265
QY 368 AAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACGATGACGCTCATCGAAGCAAA 427

```
Db 266 CGATCCGCGACTGCTGCTGGGCAAGGACGCGCGCCGACGAGAAAGCCCTCGACACCGCA 325
Qy 428 TGATCAAGCTTTGATGGGCAACCGCAAAAGTCCGCTGGGTGCAAAAGCAATCTTGGTG 487
Db 326 TGATCGAGCTGGACGGCAACGAGAAAGGCAAGCTGGGGCCCAACGCGATCTTCGGG 385
Qy 488 TTTTCATGGCTGTGCAAAAGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCGCTACA 547
Db 386 TGTCCCTGGCTGCGCCCAAGGCGCGCACAGGCAAGGGGCTACCGCTGTACGCGACA 445
Qy 548 TCGGTGACCAAAAGCAC-----ACGTTCTTCAGTTTCCAATGATGAACATCA 595
Db 446 TCGCGCACTCAACGCACTCCCGGCCAGTACTCCATGCGCGTCCCGATGATGAACATCA 505
Qy 596 TCAACGCTGGCGCTCAAGCTGACTCGGTGTTGAGTTTCAGGATTCAGGATTCATGCTCCAA 655
Db 506 TCAACGCGCGGAGCATGCGGCAACAACTGATATCCAGGAATTCATGTCAGCGCG 565
Qy 656 TCGGTGACAGACCTTCTCTGAGGCTCTCGCGCAACGCGCGGAGTCTACACGCACTGA 715
Db 566 TCGGCGCGAAGACTTTCGCGAGGCGCTGCGCATGGGCGCGAGATCTTCATCACCTCA 625
Qy 716 AGTCCGCTCATCAAGGAAAGGCTGTCCACCGGACTTTGGCGATGAGGGCGGCTTCGCTC 775
Db 626 AGGCGGTGCTGAAGGCGCGTGGCTGAACACCGCGCTCGGTGACGAAGGCGGCTTCGCGC 685
Qy 776 CTTCCGTCGCTCCACCGTGAAGCTCTTGACCTTATCGTTTGAAGCAATCGAAGGCTG 835
Db 686 CGAACCTGTGCTCAAGCAAGACGCTTGGCGGCCATCGCGAGGCGCTCGAGAAGCGG 745
Qy 836 GCTTCAACCCAGGCAAGCAATCGCTCTGCTGGAAGTTCGTTTCTCTGAGTTCTTCA 895
Db 746 GCTACAAGCTGGGCGACGACTGACCTTGGCGCTTGACTGGCGCTCCAGCGAGTTCTTCA 805
Qy 896 AGGACGCACTTACCCTTGAAGGTGGCCAGCA-----CTCGCAAGCTGAGATGSCAA 949
Db 806 AGGACGCAAGTACGACTGGAAGCGCAAGCAAGGTATTTCAGCGCGCGTTCGCGG 865
Qy 950 ACGTTTACGTGAGCTGTTGACGCGTACCAATCTGTTCCATCGATGAGGACCCACTGAGG 1009
Db 866 ACTACTGGCGCGGCTGACCCAGCGCTACCGGATCATCTCCATCGAGGACGGCATGAGC 925
Qy 1010 AAGATGACTGGAGGGTTACACCAACTCACCGCAACCATCGGCGCAAGGTTTCAGATCG 1069
Db 926 AGTCCGACTGGCGCGGCTGGAAAGGCTTGACCGACAGATCGGCGCCCAAGGTCCAAGTG 985
Qy 1070 TTGGCGACGACTTCTTCGTCACCAACCTTGAGCGCTTGAAGGAGGCGATCGCTAAGAAGG 1129
Db 986 TCGGCGACGACCTGTTCTGAGCAACCAACCAAGATCTCTCAAGGAAGGATCGAGAAGGCA 1045
Qy 1130 CTGCAACTCTGCTGTTAAGGTGAACAGATCGGTGATCCCTCAGCGAGACTTCGAGC 1189
Db 1046 TCGGCAATTCGATCTGATCAAAGTTCAACAGATCGGTTCGCTCAGCGAGACCTCGAGG 1105
Qy 1190 CTGTCGACATGGCTCACCGCGAGGCTACCTCCATGATGTCACCGTTTCGGGTGAGA 1249
Db 1106 CCATCCAGATGGGCGCGCGGCTATACCGGGGTGATCTCGCAACGCTTCGGCGGCAAA 1165
Qy 1250 CCGAGGACACCACTTCTGCTGACCTTCGAGTTGCACTCAACTGTGGCGAGATCAAGACTG 1309
Db 1166 CCGAGGACTCGACCATCGCGGCTGCGCGGTGGGTACGCGCGCGGTGATCAAGACCG 1225
Qy 1310 GTGCTCAGACGCTTCGACGCTGTGCAAGATGACACAGCTTCTCCGATCGAGCAGC 1369
Db 1226 GTTCGCTGTGCGCTCCGACCGCGTGTCAAAGTACAAACCAAGTTGCTGCGCATCGAAGAGC 1285
Qy 1370 TGTTCGCGAGCGCGGCTTACCGAGGTCGACGCGCATTCACAGCT 1417
Db 1286 AACTGGGCGCAAGGCGCGTACCGGTGTCGCGGGAATTCGCGGCT 1333
```

```
US-09-252-991A-4158/c
; Sequence 4158, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4158
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4158
```

```
Query Match 33.7%; Score 531.4; DB 4; Length 1347;
Best Local Similarity 65.0%; Pred. No. 3.6e-139;
Matches 841; Conservative 0; Mismatches 431; Indels 21; Gaps 3;
Qy 146 CCACAGTGGCTGAAATCATGCAGTATTCGCTCGCGAAATTCCTCGACTCCCGGGTAACC 205
Db 1343 CAAGATGGCAAGATCGTCGATCAAGGGCGCTGAGTCTTGGACTCCCGCGCAACC 1284
Qy 206 CAACCTGCGAGGAGAGGTTTCTCGATGACGGTTCCTCCACGGTTCGACAGGTGTTCCAT 265
Db 1283 CTACCGTTGAAGCGGAGCTGATCTCGACAAAGCATCGTCGCGAGCGCTGCGCGCTT 1224
Qy 266 CCGGCGCATCCACGGGTCACAGGCTCATGAGTTCGCTGAGTGGGCGA---TCGCT 322
Db 1223 CCGGTGCTTCCACGGTTCCCGGAGGCGCTCGAGTTCGCGATGGCGATGGCGACAAGAGCGTT 1164
Qy 323 ACCTGGCAAGGCGGTTTGAAGGCAAGTTGAAACGTCACGAAGAAATCGGCACGAGC 382
Db 1163 ACCTGGCAAGGCGGTTGCTGAAAGCGTGGCAACATCAACGCCCGATCCGCGACTGC 1104
Qy 383 TCCTGCGCTTAGAGGTGACGATCAGCGCTTCATTCGACGAAGCAATGATCAAGCTTGATG 442
Db 1103 TGCTGGCAAGGACGCGCGCGACAGAAAGCCCTCGACACGCGATGATCGAGCTGAGC 1044
Qy 443 GCACCGCAACAGTCCCGCTCGGTGCAAAAGCAATCTTGGTGTTCATGGCTGTTG 502
Db 1043 GCACCGCAACAGGCGCAAGCTGGCGCGCAACCGCATCTCTCGCGGTGTCCTGCTGCCG 984
Qy 503 CAAAGGCTGCTGATTCGCGAGGCTCCACTGTTCCGCTACATCGGTGGACCAAGC 562
Db 983 CCAAGGCGCGCGACAGGCGCAAGGGCGTACCGGTGACGCGCATATCCCGACTCAAGC 924
Qy 563 CAC-----ACGTTCTTCCAGTTCCAAATGATGAACATCATCAACGCGTGGCGCTC 610
Db 923 GCATCCCGCGCACTACTCCATCGCGTGCAGATGATGAACATCATCAACGCGCGAGC 864
Qy 611 ACCTGACTCCGGTGTGAGTTGAGTAATTCATGATTCGCTCCAAATCGGTGCGAGACCT 670
Db 863 ATCGGACAAACAGCTCGATATCCAGGAATTCATGTGTCAGCGGTTCGCGCGCAAGAACT 804
Qy 671 TCTCTGAGGCTCTCCGCAACGCGCGAGGTCTACCAACGCTACGAGTCCGCTCATCAAG 730
Db 803 TCAGGAGGCGCTGCGCATGAGGCGCGGAGATCTTCCATACCTCAAGGCGGTGCTGAGG 744
Qy 731 AAAAGGCGCTGTCCACCGGACTTTGGCGATGAGGCGCGGCTTCGCTCTCTTCGCGGTCCA 790
Db 743 CCGGTGCGCTGAACAACGCGCTGCTGAGCAAGGCGGCTTCGCGCGCAACCTGTCGCTCCA 684
Qy 791 CCGGTGAGGCTCTTGACCTTATCGTTAAGCAATTCGAGAAAGGCTGCTTCAACCCAGGCA 850
Db 683 ACGAAGACGCGCTGGCGCGCATCGCGAGGCGCTCGAAGGCGCGGTACAAGCTGGCGC 624
```


Db 1271 TGAGCCGTTCTGACCGGTTGCTAAATACACACAGCTGATTCGTATCGAAGAAGCGCTGG 1330
Qy 1376 GCGACGCGCGCTCTACGAGTCCGACGCGCATTCGCCACG 1415
Db 1331 GTGAGCAAGCGCGTTCAACGCTGTAAGAGATCAAGG 1370

RESULT 10

US-09-818-780-22
; Sequence 22, Application US/09818780
; Patent No. 6677146
; GENERAL INFORMATION:
; APPLICANT: McHenry, Charles
; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME
; FILE REFERENCE: 1794.0030004
; CURRENT APPLICATION NUMBER: US/09/818,780
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/192,736
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Thermus thermophilus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Includes 5' UTR and 3' UTR
US-09-818-780-22

Query Match 28.5%; Score 449.2; DB 4; Length 2363;
Best Local Similarity 64.3%; Pred. No. 6.6e-116;
Matches 692; Conservative 0; Mismatches 378; Indels 6; Gaps 1;

Qy 160 ATCATGCACGATTCGCTCGCGAAATTCCTGACTCCCGCGGTAAACCAACCGTCGAGCA 219
Db 1279 ATCTCGCGCTCCGCGCACGAGGTTTGGATTCCAGGGGCTTTCCACGGTAGAGCG 1338
Qy 220 GAGGTTTTCTGGATGACGTTCCACGCTGTCCAGGTGTTCATCCGGGATCCACC 279
Db 1339 GAGGTGAGCTTGAAGCGGGGCCAGGGCGGCGCCATGTGGCCCTCCGGGGCTCCACC 1398
Qy 280 GGGCTCCACGAGGCTCATGAGCTCGGTGACGGTGGCGATCGTACCTGGGCAAGGGCGTT 339
Db 1399 GGAACCAACGAGGCTTGGAGCTCAGGACGGCGGCAAGCGCTACCTGGGCAAGGGGTG 1458
Qy 340 TTGAAGCAGTTGAAACGTTCAACGAAGAATCCGGGACGAGCTCGTGGCCTAGAGCT 399
Db 1459 CGCGGGCGGTGGAGAAGCTCAACGAGCGCATCGCCCGAGCTCGTGGCATGGACGCC 1518
Qy 400 GACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGATGGCAGCGCCACAGTCC 459
Db 1519 CTGACGAGGAAGGGGTGGAGCGGGCCATGCTGGAGTGGACGCGACCCCAACAGGCC 1578
Qy 460 CGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTCGAAAGGCTGCTGCTGAT 519
Db 1579 AACCTGGAGGAACGCGCTCTCGCGTCTCCCTGGCCGTGGCCGGCGGCGGCCAG 1638
Qy 520 TCCGAGGCTCCACATGTTCCGCTACATCGGTGGACCAACGACACAGTTCCTTCCAGTT 579
Db 1639 GCCCTGGGCTGCCCTTTTACCGCTACTCTGGCGGGGTTCAGGGGGTCACTTCGCCGTG 1698
Qy 580 CCAATGATGAACATCATCAAGGTGGCGCTCACGCTGATCCCGTGTGAGTTTCAGAA 639
Db 1699 CCCCTCATGACGTCATCAACGGGGGAAGCACGCCGACCAACCGGGTGGACTTCAGAG 1758
Qy 640 TTATGATCGCTCCAATCGGTGAGAGACCTTCTCGAGGCTCTCCGCAAGCGGCGGAG 699
Db 1759 TTATGCTGTGCGCGGGCGGGAAGCTTTCGCGAGGCTTGAGATCGGGGCCAG 1818
Qy 700 GTCTACACGACCTGAGATCCGTATCAAGGAAAGGGCTGTCCACCGGACTTGGCGAT 759

Db 1819 GTCTTCCACACCCCTCAAGCGCGTCTCAAGAGAAGGGCTACAGCAACAAAGTGGGGAC 1878
Qy 760 GAGCGGGCTTCGCTCTCTCCGTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAG 819
Db 1879 GAGGAGGCTTCGCCCGGACCTCAGGAGCAAGAGAGCGGTGGAGCTTTGCTCCTC 1938
Qy 820 GCAATCGAAGAGCTGGCTTCAACCCAGGCAAGCATCGCTCTTGTCTGGAGCTTGTCT 879
Db 1939 GCCATTGAGCGGGCGGGGTACACCCCGGGCCAGAGGTCTCCCTGGCCCTGGACCCGCC 1998
Qy 880 TCCTCTGAGTTCTTCAAGGAGGCACTTACACTTCGAAGTG-----GCCAGCACTCC 933
Db 1999 ACAGCGAGCTTTACCGGGACGGGAAGTACCACTGGAAGGGAGGCAAGGTCTCTCC 2058
Qy 934 GCAGCTGAGATGCAAAACGTTTACGCTGAGCTGTTGACGCGGTACCCCAATCGTCTCCATC 993
Db 2059 TCGAGGAGATGTTGGCTTCTCGGAGGCTTGGTGGAGAGTACCCCATCCGCTCCATT 2118
Qy 994 GAGGACCCACTCGAGGAAGATGATGGGAGGGTTACACCAACCTCACCGCAACCATCGGC 1053
Db 2119 GAGGACGGCTTCCGAGGAGCACTGGGAGGGGTGGCGGCTTCTCACCGAGCGCTGGG 2178
Qy 1054 GACAAGTTTCAAGTCTGGGAGCACTTCTTCTCAACCAACCTTGAGCGCTGAAAGAG 1113
Db 2179 GGAAGGTTCCAGCTCGTGGGGGACGACCTTCTCTCAACCAACCGGAAAGGCTCCGGCG 2238
Qy 1114 GGCATCGCTAAGAAAGGCTGCCTTCCATCTCGTTAAGGTGAACAGATCGGTACCTC 1173
Db 2239 GGAATTGAGCGGGGGTGGCCAAAGCCATCTGTCGTAAGGTGAACAGATCGGAGCCCTC 2298
Qy 1174 ACCGAGACCTTCGACGCTGTCGATGCTGTCACCGCGAGGCTACACCTCCATGAT 1229
Db 2299 TCGGAAACCTTCGAGGCTCCGCTGGCCCGCTCGGGGTACAGGGCGGTGAT 2354

RESULT 11

US-08-781-802-9
; Sequence 9, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIKENS, John
; APPLICANT: FOSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,802
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,078
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704


```

; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,995
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3545 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1397..2905
; OTHER INFORMATION: /note= "E019 sequence of longest
; OTHER INFORMATION: open reading frame; upstream untranslated region not exact"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1397..2905
; US-08-781-802-9

```

```

Query Match      27.9%; Score 440.8; DB 2; Length 3545;
Best Local Similarity 62.7%; Pred. No. 1.9e-113;
Matches 742; Conservative 0; Mismatches 427; Indels 15; Gaps 3;

Qy 229 CTGATGACGGTCCCAACGGTGTGCGAGGTGTTCCATCCGCGCATCCACCGCGTCCAC 288
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 289 GAGCTCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 345
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 99 GAAGCGTGGTAAATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 158
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 346 GCAGTTGAAACGTCACAGGAATCGGCGAGCTCGTGGCTTAGAGGCTGACGAT 405
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 159 GCGGTTGAGAACGTCACGAAGTGTGTTGCTCGGGAATCATCGCTTAGAAGTACGTAT 218
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 406 CAGGCGCTCATCGACGAAGCAATGATCAAGTTGATGGCACCGCCAAACAAAGTCCCGCTG 465
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 219 CAAGTGGCGATCGACCGCGGTTGATTGAATTTGACGGCACGGAACAAAGGAAAGCTT 278
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 466 GTGCAACACCAATCCCTTGGTGTTCATCGCTGTTCGAAGCTGCTGATTCGCA 525
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 279 GGGCGAATGCTATTTAGCGGTGTCGCTCGCGTGTGCTCGCGTGGGCTGATGAGCTT 338
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 526 GGCTTCCACTGTTCCGCTATACATGGTGGACCAACACACAGTTCCTCCAGTTCCAATG 585
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 339 GGCTTGGGTTGACCAATACTTGGCGGCTTTAAACGCTAAACGCTGCTGTACCGATG 398
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 586 ATGAACATCATCAACGGTGGCGCTCAGCTGACTCCGGTGTGACGTTTCAGGAATTCATG 645
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 399 ATGAACATTTTAAACGGCGCGCGCATGCGGACAAACAACTTGACATTTCAAGAATTCATG 458
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 646 ATCGCTCCAATCGTGAGAGACTTCTCTGAGCTCTCGCAACGCGCGGAGGTTCTAC 705
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 459 ATATGCGGTCGGTGGGAAAGCTTCCGTGAAGCGCTGCGCATGGGTGCAAAATTTTC 518
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 706 CAGCACTGAAGTCCGTCATCAAGGAAAGGGCTGTCCACCGGACTTGGCGATGAGGGC 765
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 519 CATAGCTTAAAGCTGTGTTTAAAGCGAAAGGCTTACAAACACGCTGTGCGGTGACGAAGC 578
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 766 GGCTTCGCTCTTCGTCGCGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGAATC 825
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 579 GGATTTGCTCCGAACCTTAAATCGAACGAAGAGCGCTGCAACACGATCATTTGAAGCATC 638
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 826 GAGAAGCTGGCTTACCCCGAGGAGACATCGCTCTTGTCTGTGACGTTGCTTCTCT 885
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 639 GAAAAAGCCGCTACAAACGAGGCAACAGTGTGCTGTATGACGTTGCTTCTGTCG 698
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 886 GAGTTCTTCAAG-----GAGGCACTTACCACTT-----CGAAGTGGCAGCACTCC 933
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 699 GAGCTGTACAAAGAGATGCAATATCATTTGGAAGGCGAAGGCGTGGTCAAAACA 758
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 934 GCAGCTGAGATGGCAACGTTTACGCTGAGCTGTGACGCGTACCAATCGTCTCCATC 993
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 759 TCAGAGAAATGTTGCTTGGTATGAAGAGCTTGTGCGAAATATCGATCATCTCGATC 818
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 994 GAGGACCCACTGCAGGAAGATGATGGAGGGTTACCAACCTCACCAGCAACATCGGC 1053
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 819 GAAGACGGACTTGCAGAAATGATGGAAGGCGCATAACTGCTTACTGAGCGCTTGGC 878
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1054 GACAAGTTTCAGATCGTTGCGGACGACTTCTGTCACCAACCTGAGCGCTGAAGGAG 1113
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 879 CACAAAGTGCAGCTCGTGGTGACGACTTGTGTTGTAACGAACAGAAAACTGGCGGAA 938
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1114 GGCATCGCTAAGAAAGCTGCCAATCCATCTCTGTTTAAAGTGAACAGATCGGTACCTC 1173
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 939 GGCATTGAAAAAGGCGTGGCAACTCGATTTTAAATTAAGTGAACCAATCGGTACACTG 998
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1174 ACCGAGACCTTCGACGCTGCGACATGGCTCACCAGCGAGGCTACACCTCCATGATGTC 1233
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 999 ACGAAACGTTTCGATGCTTGGATGGCAACCGCGCGGTACACGCGGTTGTGTGTCG 1058
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1234 CACGTTCCGCTGAGACCGAGGACACCATTTGCTGACTCGCAGTTGCACTCAACTGT 1293
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1059 CACGTTCCGTTGAAAGGAGACAGACAGATTCGCGATATCGCTGCGCAACAAACGCT 1118
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1294 GGCAGATCAAGACTGGTGTCCAGCACGTTCCGACCGTGTCCAAAGTACAAACAGCTT 1353
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1354 CTCGCGATCGAGCAGCTGCTTGGCAGCGCGGGGTCTACGAGG 1397
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1179 CTCGCGATTGAAGACGAACCTTGGCCACAGGCTATTTACCAAGG 1222
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
US-09-583-110-1156
; Sequence 1156, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1156
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1156

```

```

Query Match      27.4%; Score 431.6; DB 4; Length 1305;
Best Local Similarity 59.8%; Pred. No. 4.3e-111;
Matches 775; Conservative 0; Mismatches 494; Indels 27; Gaps 2;

Qy 160 ATCATGACGATATTCGCTCGCAAAATTCGACTCCCGGGTAAACCAACCGTCGAGCA 219
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 10 ATTACTGATGTTTACGCTCGCAAGTCTTAGACTCAGCGGTAACCCCAACACTTGAAGTA 69
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 220 GAGGTTTCTCGATGACGTTCCACGGTGTGCGAGGTGTTCATCCGCGCATCCACC 279
DB 70 GAAGTTTACACTGAACTCAGGTGCTTTCCGACGTGTATGTTCCATCAGGAGCTTCTACT 129
QY 280 GGGCTCCACGAGGCTCATGAGCTCGTGACGGTGGCA---TCGCTACCTGGGCAAGGC 336
DB 130 GGTGAACACCAAGCAGTTGAACCTTCGCGACGGTGACAAATCTCGTTACGGTGGTCTTGGT 189
QY 337 GTTTTGAAGCGAGTTGAAACGTCACGAAGAAATCGCGACGAGCTCGCTGGCCTAGAG 396
DB 190 ACACAAAAGCTGTGACAGCTAAACAACATCATCTGCTGAAGCTATCATTTGGCTAGCAT 249
QY 397 GCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTGTATGGCACCGCAACAAG 456
DB 250 GTACGCTGATCAACAAGCTATTGACCGTGTATGATCGCACTTGACGGTACTCTCTAACAA 309
QY 457 TCCGCGCTGGTGAACAGCAATCCTTTGGTGTTCATGCTGTGTGCAAAAGGCTGTGCT 516
DB 310 GGTAAATTTGGGTGGAAATGCAATCCTCGGTGTGTCTATCGCTGTAGCTCGTGTGCTGCT 369
QY 517 GATTCCGACGCTCCCACTGTTCCGCTACATCGGTGACCAACGACACACGCTTTTCCA 576
DB 370 GACTACCTTGAATCCCACTTTACAGCTACTTGTGTGATTCACACTAAAGTTCTTCCA 429
QY 577 GTTCCAAATGATGAACATCATCAACGCTGGGCTCACGCTGACTCCGCTGTGTGAGCTTCAG 636
DB 430 ACTCCAAATGATGAACATCATCAACGCTGGTGTCTCACTCTGACGCTCCCAATCGCTTCCA 489
QY 637 GAATTCATGATCGTCCAAATCGGTGACAGACTTCTCTGAGGCTCTCCGCAACGGCGG 696
DB 490 GAGTTCTATGATCTTGCAGATGTTGGTGCACCAACATTTAAAGAAAGCCCTTCGTTACGGTGT 549
QY 697 GAGTCTACCAACGACATGAAGTGGTCTCATCAAGAAAGGCGCTGTCCACGGACTTGGC 756
DB 550 GAAATCTTCCAGCGCTTTAGAAATCCTTAATACGCTGTTTGAAGCTGCGGTAGGT 609
QY 757 GATGAGGCGGCTTCTGCTCTCTCCGCTGCCAGCTGAGGCTCTTGACCTTATCGTT 816
DB 610 GACGAGGTGATTCGCTCTCTCGTTTCGAGGAACTGAAGTGTGTGAACATCTCTT 669
QY 817 AAGCAATCAGAAAGGCTGCTTCAACCCAGGCAAGACATCGCTCTTGTCTGAGAGTT 876
DB 670 GCTGCGATTGAAGCTGTGATATGATACAGGTAAGACGCTATTATTGGAATTTGACTGT 729
QY 877 GCTTCTCTGAGTCTTCAAGGACGACCTACCACTTCA----- 917
DB 730 GCTTCTACGAAATTTCTACGATAAGAAACGTAAGTTTACGACTACACTAAATTTGAAGT 789
QY 918 -----AGTGGCCAGCACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGAC 972
DB 790 GAAGGTGCTGCTGTCTGATCATCTGACAGAACAAATCGACTACCTTGAAGAAATTTGTTAAC 849
QY 973 GCGTACCCTAATCGTCTCCATCGAGGACCCACTGACAGAAAGATGACTGGGAGGTTACACC 1032
DB 850 AAATACCAATCATCACTATTGAAGATGGTATGATGAAACGACTGGGATGGTTGAAA 909
QY 1033 AACCTACCCCAACCATCGGCAAGGTTTCAGATCGTTGGCGACGACTTCTTGGTCAAC 1092
DB 910 GCTTCTAAGAGCTCTTGTGTGAAGAAAGTAACTTGTGGTGAACGACTTCTTCTGTAACA 969
QY 1093 AACCTCAGGCGCTTGAAGGAGGAGCTGCTTGAAGGCTGCAACTCCATCTCTGTTTAAAG 1152
DB 970 AACACTGACTACCTTGCAGCTGTATCAAGAGGCTGCTGCTAACTCAATCTTATCAAA 1029
QY 1153 GTGAACAGATCGGTACCCCTCACCGAGACCTTTCGACGCTGTGACATGGCTACCGGCA 1212
DB 1030 GTTAAACCAATCGGTACTCTTACTGAAACCTTTTGAAGCTATCGAAATGGCTTAAAGAGCT 1089
QY 1213 GGCTACACCTCCATGATGTCCACCGGTTCCGGTGAGACCCAGGACACCACTTGTGCTGAC 1272
DB 1090 GGTTACACTGCTGTGTATCACACCGGTTTCAAGTGAAGTAACTGAAGTTCACCAATCGCTGAT 1149

QY 1273 CTCGCACTTGCACTCAACTGTGGCAGATCAAGACTGGTCTCCAGCACGCTTCCGACCGT 1332
DB 1150 ATTGCACTTGCACTCAACTAAGCGAGGACAAATCAAGACTGGTTCATCTTTCCGCTACAGACCGC 1209
QY 1333 GTCCGAAAGTACAAACAGCTTCTCCGATCGACGAGCTGCTTGGCAGACGCGGCTGTAC 1392
DB 1210 ATCGCTAAATACAAACAAATTTGCTTCGATTCGAAAGACCAACTTGGTGAAGTGTGATAT 1269
QY 1393 GCAAGTGGCAGCGCACTTCCACGCTTTTCAGGGCTAA 1428
DB 1270 CGTGGATTGAATCATTTCTACAACCTTAAAAATAA 1305

RESULT 13
US-08-961-527-45/c
; Sequence 45, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-45

Query Match 27.3%; Score 430.4; DB 3; Length 11384;
Best Local Similarity 59.4%; Pred. No. 2,9e-110;
Matches 781; Conservative 0; Mismatches 506; Indels 27; Gaps 2;
QY 142 GAGGCCACAGTGGCTGAAATCATGACAGTATTTCGCTCGGAAATTCCTCGACTCCCGCGT 201
DB 3064 GAGTAAAAATGTCATTAATTAATCTGATGTTTACGCTCGCGAAATCTCTAGACTCACGCGT 3005
QY 202 AACCCAAACGTCGAGGACAGGTTTTCCTGGATGACGGTTCACCGGTTCGCGAGGTGT 261
DB 3004 AACCCAAACCTTGAAGTAGAAGTTTACACTGAATCAGGTGCTTTCGGAACGTTGATGGT 2945
QY 262 CCATCCGCGCATCCACCGCGCTCCAGAGGCTCATGAGCTGCGTGCAGGTGCGGAGT 318
DB 2944 CCATCAGGAGCTTCTACTGTGTGAACGACGAGGTTGAATCTTCGCGCGGTGACAAATCT 2885
QY 319 CGCTACTCGGCGAAGGCGCTTTTGAAGGCGAGTGTGAAAAACGTCACGAAGAAATCGCGCAC 378

Db 2884 CGTTACGGTGGCTCTTGGTACACAAAAAGCTGTGTGACAAAGTAAACAAATCAATTGCTGAA 2825
Qy 379 GAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAGCAAGCAATGATCAAGCTT 438
Db 2824 GCTATCATTTGGCTACGATGACGTGATCAACAAGCTATTGACCGTGTATGATCGCACTT 2765
Qy 439 GATGGCACCGCCAAACAAGTCCGCTCGGTCGAAACGCAATCCCTTGGTGTGTTCCATGGCT 498
Db 2764 GACGGTACTCTTAACAAGGTAATTTGGTGGATGCAATCTCGTGTGCTATCGCT 2705
Qy 499 GTTGCAAAAGGTGCTGCTGATTCGCGAGGCTCCACTTTCGCTACATCGGFGACCA 558
Db 2704 GTAGCTCGTGTGCTGCTGACTACTTTCGAAATCCCACTTACAGCTATCTTTGGTGGATTC 2645
Qy 559 AACGCACACGTTCTTCAGTTCGAATGATGACATCATCAACGGTGGCGCTCAGCTGAC 618
Db 2644 AACACTAAGTTCCTCAACTCCCAATGATGAACATCATCAACGGTGGTTCCTCACTCTGAC 2585
Qy 619 TCCGGTGTTCAGCTTCAGGAATTCATGATCGCTCCAATCGGTGCGAGACCTTCTCTGAG 678
Db 2584 GCTCCAATCGCTTTCAGAGTTCATGATCTTGCCAGTTGGTGGCCAAATTTAAAGAA 2525
Qy 679 GCTTCGCGAACGCGCGGAGGTCTACACGCACTGAAGTCCGTCATCAAGGAAAGGCG 738
Db 2524 GCCCTTCGTTACGGTGTGCTGAAATCTTCAGGCTCTTAAGAAAAATCTCTAAATCAGGTGT 2465
Qy 739 CTGTCCACCGGACTTGGCGGATGAGGCGGCTTCGCTCTTCCGTCGGCTCCACCGGTGAG 798
Db 2464 TTGGAACCTGCGTAGTGAGAGGTGGATTCGCTCTCGTTTCGAAGGAATCGAAGAT 2405
Qy 799 GCTCTTACCTTATCGTTAAGGCAATCGAGAAGCTGGCTTCACCCAGGCAAGGACATC 858
Db 2404 GGTGTGAAACTATCTCTTCTGGATTAAGCTGCTGGATATGTACCAGGTAAAGAGTA 2345
Qy 859 GCTCTTCTCTGAGCGTGTCTTCTGAGTTCCTTAAGGACGGCACTACCACTTCGAA 918
Db 2344 TTTATCGGATTTGACTGTGCTGCTTATCAGAATTTACGATAAGAAAGTAAAGTTTACGAC 2285
Qy 919 -----GGTGGCGAGCACTCCGACGTGAGATGGCAACGTT 954
Db 2284 TACATAAATTTGAAGTGAAGGTGCTGCTGTTGCTGATCTCGAGAACAAATCGACTAC 2225
Qy 955 TAGCTGAGCTCGTTGACGGTACCCAAATCGTTCCTCAAGGACCACTCGAGGAAGAT 1014
Db 2224 CTTGAAGAAATGGTTAAACAATAACCCAAATCATCACTATTGAAGATGATGATGAAAC 2165
Qy 1015 GACTGGGAGGTTTACACCAACCTCACCGCAACCATCGGCGACAAGGTTGAGATCGTTGGC 1074
Db 2164 GACTGGGATGGTTGGAAGCTCTTACTGAAACGCTTTGTTGTAAGAAAGTACAACTTGTGTT 2105
Qy 1075 GAGGACTTCTTCTGTCACCAACCTGAGCGCTGAAAGGAGGCACTCGCTAAGAGGCTGCC 1134
Db 2104 GAGGACTTCTTCTGTAACAACACTGACTACTCTTGACGTGGTATCCAAAGAGTGTGCT 2045
Qy 1135 AACTCCATCTCGTTAAGGTGAACCGATCGGTACCTCAACCGAGACCTTCGACGCTGTC 1194
Db 2044 AACTCAATCTTATCAAAAGTTAACCAATCCGTACTCTTACTGAAACTTTTGAAGCTATC 1985
Qy 1195 GACATGGCTCACCGCGAGCTACACCTCCATGATGTCCACCGTTCGCGTGAGACGAG 1254
Db 1984 GAAATGGCTTAAAGAGCTGTTTACCTGCTGTTGATCAACACCGTTTCAGGTGAACCTGAA 1925
Qy 1255 GACACCACTATGCTGACCTCGAGATTGCACTCAACTGTGGCGAGATCAAGACTGGTGTCT 1314
Db 1924 GATTCAACATCGCTGATATTGAGTTGCACTTAAACGAGGACAAATCAAGACTGGTTCA 1865
Qy 1315 CAGACGTTTCGACCGGTGCGAAAGTAAACACAGCTTCTCCGATTCGAGCAGCTGCTT 1374
Db 1864 CTTTCAGGTACAGACCGCATCGCTTAAATACAAACCAATGCTTCTGATCGAAGACCAACTT 1805
Qy 1375 GCGGAGCGCGCTCTACGAGGTCGAGCGGATTCCTCCACGCTTTCAGGCTTAA 1428
Db 1804 GGTGAAGTAGCTGAATATCGTGAATTGAAATCATCTTACAACTTAAAAAATAA 1751

RESULT 14
US-09-107-532A-156
; Sequence 156, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1196
; SEQUENCE DESCRIPTION: SEQ ID NO: 156:
US-09-107-532A-156

Query Match 27.1%; Score 428; DB 4; Length 1196;
Best Local Similarity 61.4%; Pred. No. 4.3e-110;
Matches 730; Conservative 0; Mismatches 440; Indels 18; Gaps 2;
Qy 259 GTTCCATCGGCGCATCCACCGCGTCCACGAGGCTCATGAGCTGCGTGACCGTGGC--- 315
Db 6 GTTCTTCCAGTGTCTTCAACTGGTGAATACGAAGCAGTTGAATTACGTGACCGAGACAA 65
Qy 316 GATCGCTACTGGGCGAGGCGGTTTTGAAGGCGAGTTGAAAACGTCAACGAAGAAATCGGC 375
Db 66 GCTCGTTATGGCGGTAAAGGTGTAATAAGCTGTGACAAACGTAAACCAATCATTTGCT 125
Qy 376 GACGAGCTCGCTGGCGCTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAG 435
Db 126 GAAGCAATCATTTGGCTACGATGTTCTGACCAAAATGGCTATCGACAAAAGCTATGCTCT 195
Qy 436 CTTGATGGCAACCGCCAAAGTCCCGCTGGGTGCAACGCAATCCTCGTGTTCATG 495

Qy	1081	TTCTTCGTACCAACCTGAGCGCTGAAGAGGGCATCGCTAAGAGGCTGCCAACTCC	1140
Db	393	CTGTTGCTAAACCAACACCAAGATCCTGAAAAGAGGCATCGAAAGAGGCATCGCTAACTCC	334
Qy	1141	ATCCTGGTTAAGGTGAACCAAGATCGGTACCTCACCAGAGACCTTCGACGCTGTCGACATG	1200
Db	333	ATCCTGATCAAGTTCAACCAAGATCGGTCTCTGACCGAAACTCTGGCTGCTATCAAGATG	274
Qy	1201	GCTCACCGCGAGGCTACACCTCCATGATGCCACCGTTCCGGTGAAGACCGAGGACACC	1260
Db	273	GCGAAAGACGCTGGCTACACCGCTGTTATCTCTCACCGTTCTGGCGAAACTGAAGACGCT	214
Qy	1261	ACCATTCCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA	1320
Db	213	ACCATTCGCCACCTGGCTGTGTGTACCGCTGCAAGGCCAGATCAAAACTGGTTCATGAGC	154
Qy	1321	CGTTCGACCGTGTGCAAAAGTACAACAGCTTCTCCGCATCGAGCAGCTGCTTGGCGAC	1380
Db	153	CGTTCGACCGGTTGCTAATACACCAAGCTGATTCGTATCGAAGAGCGCTGGGTGAG	94
Qy	1381	GCCGGGCTCTACGAGGTCGAGCGCATTCCTCAG	1415
Db	93	CAAGCGCGTTCAACGGTCGTAAAGAGATCAAAGG	59

Search completed: September 28, 2005, 23:33:54
 Job time : 293.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2005, 20:45:32 ; Search time 4241.5 Seconds
(without alignments)
2568.796 Million cell updates/sec

Title: US-10-728-947-3
Perfect score: 1578
Sequence: 1 ggctgggatatggtagttt.....ctcaagcagggaactgctt 1578

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7442561 seqs, 3452328358 residues

Total number of hits satisfying chosen parameters: 14885122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1578	100.0	1578	9	US-09-860-768-3
2	1578	100.0	1578	20	US-10-728-947-3
3	1576.4	99.9	1578	9	US-09-860-768-1
4	1576.4	99.9	1578	20	US-10-728-947-1
5	1576.4	99.9	3309400	9	US-09-738-626-1
6	1401.8	88.8	1405	21	US-10-494-836-7
7	1394.8	88.4	1398	19	US-10-781-014-71

8	1273.4	80.7	1275	9	US-09-738-626-1085	Sequence 1085, Ap
9	907	47.5	1275	17	US-10-282-122A-17724	Sequence 17724, A
c	696.8	44.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
11	695.4	44.1	1284	15	US-10-156-761-3520	Sequence 3520, Ap
12	678.2	43.0	1269	17	US-10-369-493-32027	Sequence 32027, A
13	647	41.0	1287	17	US-10-282-122A-25949	Sequence 25949, A
14	632.6	40.1	45191	15	US-10-080-170-649	Sequence 649, App
15	632.6	40.1	45191	19	US-10-080-170-649	Sequence 649, App
16	632.6	40.1	45191	19	US-10-468-356-649	Sequence 649, App
17	632.4	40.1	1317	17	US-10-282-122A-26512	Sequence 26512, A
18	626.2	39.7	1290	17	US-10-282-122A-28306	Sequence 28306, A
19	583.4	37.0	1344	17	US-10-282-122A-27498	Sequence 27498, A
c	583	36.9	2256646	19	US-10-470-565-1	Sequence 1, Appli
21	580.4	36.8	1269	17	US-10-369-493-34244	Sequence 34244, A
22	577	36.6	1269	17	US-10-369-493-35383	Sequence 35383, A
23	575.6	36.5	1281	17	US-10-369-493-40598	Sequence 40598, A
24	572.4	36.3	1275	17	US-10-369-493-35506	Sequence 35506, A
25	569	36.1	1260	17	US-10-369-493-38426	Sequence 38426, A
26	569	36.1	1269	17	US-10-369-493-38909	Sequence 38909, A
27	568.2	36.0	1278	17	US-10-369-493-41573	Sequence 41573, A
28	567.4	36.0	1272	17	US-10-369-493-44349	Sequence 44349, A
29	561.8	35.6	1272	17	US-10-369-493-31449	Sequence 31449, A
30	558.6	35.4	1278	17	US-10-369-493-45417	Sequence 45417, A
31	550.4	34.9	1290	17	US-10-369-493-43178	Sequence 43178, A
32	542	34.3	1281	17	US-10-282-122A-11620	Sequence 11620, A
33	530.2	33.6	1290	9	US-09-815-242-7834	Sequence 7834, Ap
34	530.2	33.6	1290	17	US-10-282-122A-30326	Sequence 30326, A
35	529.2	33.5	1281	17	US-10-369-493-44693	Sequence 44693, A
36	528	33.5	1287	17	US-10-369-493-32181	Sequence 32181, A
37	522.8	33.1	1278	17	US-10-369-493-28375	Sequence 28375, A
38	522.8	33.1	1281	17	US-10-369-493-31134	Sequence 31134, A
39	522.8	33.1	1281	17	US-10-282-122A-13098	Sequence 13098, A
40	513.6	32.5	1404	17	US-10-282-122A-33598	Sequence 33598, A
41	513.2	32.5	1269	17	US-10-369-493-47120	Sequence 47120, A
42	512	32.4	1293	17	US-10-369-493-46986	Sequence 46986, A
43	507.8	32.2	1293	17	US-10-369-493-37536	Sequence 37536, A
44	506	32.1	1293	9	US-09-974-300-664	Sequence 664, App
45	503.2	31.9	1293	17	US-10-282-122A-9983	Sequence 9983, Ap

ALIGNMENTS

RESULT 1

US-09-860-768-3
; Sequence 3, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-3

Query Match 100.0%; Score 1578; DB 9; Length 1578;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAAATTCACCTGATTCATCGAAACAAGA 60

Db 1 GGCTGGGGATATGGGTAGTTTTCGCCCACTAAATTCACCTGATTCATCGAAACAAGA 60


```
QY 61 TTCTGTCACAAATTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTCTAG 120
Db 61 TTCTGTCACAAATTGGGTAGACGTGATTGAAGACATTTTGATCAGTGAATAATTCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGACAGTATTGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGACAGTATTGCTCGC 180
QY 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGCGAGAGTTTCTTGGATGACGGT 240
Db 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGCGAGAGTTTCTTGGATGACGGT 240
QY 241 TCCACAGGTTCGACAGTGTTCATCCGCGCATCCACCGCGTCCACAGGCTCATGAG 300
Db 241 TCCACAGGTTCGACAGTGTTCATCCGCGCATCCACCGCGTCCACAGGCTCATGAG 300
QY 301 CTGCGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGTTGAAAACGTC 360
Db 301 CTGCGTGACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGTTGAAAACGTC 360
QY 361 AACGAGAAATCGCGCAGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCCTCATCGAC 420
Db 361 AACGAGAAATCGCGCAGAGCTCGCTGGCTTAGAGGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGATGGCACCAGCAAGTCCCGCTGGGTGCAAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGATGGCACCAGCAAGTCCCGCTGGGTGCAAAACGCAATC 480
QY 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTC 540
Db 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTC 540
QY 541 CGCTACATCGGTGACCAAGGACACAGCTTCTCCAGTTCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGACCAAGGACACAGCTTCTCCAGTTCCAATGATGAACATCATCAAC 600
QY 601 GGTGGGCTCAGCTGACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGT 660
Db 601 GGTGGGCTCAGCTGACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGACCGCGCGAGGCTTACACGCACTGAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGACCGCGCGAGGCTTACACGCACTGAGTCC 720
QY 721 GTCATCAAGGAAAGGGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGGAAAGGGCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCCTTCC 780
QY 781 GTGGCTCCACCGTGAAGGCTTTGACCTTATCGTTAAGCAATCGAAGGCTGGCTTC 840
Db 781 GTGGCTCCACCGTGAAGGCTTTGACCTTATCGTTAAGCAATCGAAGGCTGGCTTC 840
QY 841 ACCCAGGCAAGGACATCGCTTTGGCTTCGAGGTTCTTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTTTGGCTTCGAGGTTCTTCTCTGAGTTCTTCAAGGAC 900
QY 901 GGCACCTACCACTCGAAGGTGGCGACACTCCGACGCTGAGATGGCAAAACGTTACGCT 960
Db 901 GGCACCTACCACTCGAAGGTGGCGACACTCCGACGCTGAGATGGCAAAACGTTACGCT 960
QY 961 GAGCTCGTTGACGGTACCCAAATCGTCTTCGAGGTTCTTCTCTGAGTTCTTCAAGGAC 1020
Db 961 GAGCTCGTTGACGGTACCCAAATCGTCTTCGAGGTTCTTCTCTGAGTTCTTCAAGGAC 1020
QY 1021 GAGGTTTACCAACCTCACCGCAACCATCGGACAAAGGTTTCAGATCGTTGGCGACGAC 1080
Db 1021 GAGGTTTACCAACCTCACCGCAACCATCGGACAAAGGTTTCAGATCGTTGGCGACGAC 1080
QY 1081 TTCTTCGTCACCAACCTGAGCGCTGAAGGAGGCTCGCTAAGAGGCTGCGCACTCC 1140
Db 1081 TTCTTCGTCACCAACCTGAGCGCTGAAGGAGGCTCGCTAAGAGGCTGCGCACTCC 1140
```

```
QY 1141 ATCTCGTTAAAGTGAAACAGATCGGTACCTCAACGAGACCTTTCGACGCTGTCGACATG 1200
Db 1141 ATCTCGTTAAAGTGAAACAGATCGGTACCTCAACGAGACCTTTCGACGCTGTCGACATG 1200
QY 1201 GCTCACCGCGCAGGCTTACACCTCCATGATGTCCCACCGTTCGGTGAGACCGAGGACAC 1260
Db 1201 GCTCACCGCGCAGGCTTACACCTCCATGATGTCCCACCGTTCGGTGAGACCGAGGACAC 1260
QY 1261 ACCATTGCTGACCTCCAGTTGACCTCAACTGTGGCAGATCAAGACTGGTGTCTCAGCA 1320
Db 1261 ACCATTGCTGACCTCCAGTTGACCTCAACTGTGGCAGATCAAGACTGGTGTCTCAGCA 1320
QY 1321 CGTTCCGACCGTGTCCGAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGTGTCCGAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
QY 1381 GCGGCGCTCTACGACGGTCCGACGCAATTCACACGCTTTCAGGGCTTAAATAAAGGCGTT 1440
Db 1381 GCGGCGCTCTACGACGGTCCGACGCAATTCACACGCTTTCAGGGCTTAAATAAAGGCGTT 1440
QY 1441 TTCGACGCGCGGTAACTCAAGTTGCCGGGCTGCTTGCCTTACTACTGTTACTGTTGT 1500
Db 1441 TTCGACGCGCGGTAACTCAAGTTGCCGGGCTGCTTGCCTTACTACTGTTACTGTTGT 1500
QY 1501 GACTATGATCGAGGATTATGGCAAGCAGAGAAACTCATAAAGGCTTGTTCCTGTCT 1560
Db 1501 GACTATGATCGAGGATTATGGCAAGCAGAGAAACTCATAAAGGCTTGTTCCTGTCT 1560
QY 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578
```

RESULT 2

```
US-10-728-947-3
; Sequence 3, Application US/10728947
; Publication No. US20040220394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10/728,947
; CURRENT FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-3
```

```
Query Match 100.0%; Score 1578; DB 20; Length 1578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAAATTTCACTGATTCGCTCATCGAAACAAGA 60
Db 1 GGCTGGGGATATGGGTAGTTTTCGCCACTAAATTTCACTGATTCGCTCATCGAAACAAGA 60
QY 61 TTCTGTCACAAATTTGGGTGTAGACGTTTGAAGACATTTGATCAGTGAATAATTCTAG 120
Db 61 TTCTGTCACAAATTTGGGTGTAGACGTTTGAAGACATTTGATCAGTGAATAATTCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGACAGTATTGCTCGC 180
Db 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGACAGTATTGCTCGC 180
```

```
QY 181 GAAATTCTCGACTCCCGCGGTAAACCAACCGTTCGAGCGAGAGGTTTTCTTGGATGACGGT 240
DB 181 GAAATTCTCGACTCCCGCGGTAAACCAACCGTTCGAGCGAGAGGTTTTCTTGGATGACGGT 240
QY 241 TCCGACGGTTCGAGGTGTTTCATCCGGGCGCATCCACCGGCGTTCACGAGGCTCATGAG 300
DB 241 TCCGACGGTTCGAGGTGTTTCATCCGGGCGCATCCACCGGCGTTCACGAGGCTCATGAG 300
QY 301 CTGGCTGACGGTTCGAGGTGTTTCATCCGGGCGCATCCACCGGCGTTCACGAGGCTCATGAG 360
DB 301 CTGGCTGACGGTTCGAGGTGTTTCATCCGGGCGCATCCACCGGCGTTCACGAGGCTCATGAG 360
QY 361 AACGAGAAATTCGCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
DB 361 AACGAGAAATTCGCGCGACGAGCTCGCTGGCTAGAGGCTGACGATCAGCGCCTCATCGAC 420
QY 421 GAAGCAATGATCAAGCTTGTATGGACCGCCAAACGAGTCCCGGCTGGGTGCAACGCAATC 480
DB 421 GAAGCAATGATCAAGCTTGTATGGACCGCCAAACGAGTCCCGGCTGGGTGCAACGCAATC 480
QY 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
DB 481 CTTGGTGTTCATGGCTGTTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTTC 540
QY 541 CGCTACATCGGTGGACCAAAACGACACGCTTCTTCCAGTTCCAATGATGAACATCATCAAC 600
DB 541 CGCTACATCGGTGGACCAAAACGACACGCTTCTTCCAGTTCCAATGATGAACATCATCAAC 600
QY 601 GGTGGGCTCAGCTGATCGGTGTTGAGTTCAGGAATTCATGATCGTTCGAATCGGT 660
DB 601 GGTGGGCTCAGCTGATCGGTGTTGAGTTCAGGAATTCATGATCGTTCGAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGTCTACACGCACTGAGTCC 720
DB 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGCGCGGAGTCTACACGCACTGAGTCC 720
QY 721 GTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
DB 721 GTCATCAAGGAAAGGCGCTGTCCACCGGACTTGGCGATGAGGCGGCTTCGCTCCTTCC 780
QY 781 GTGGCTTCCACCGGTGAGGCTTGTGACCTTATCGTTAAGGCAATCGAGAAAGGCTGCTTC 840
DB 781 GTGGCTTCCACCGGTGAGGCTTGTGACCTTATCGTTAAGGCAATCGAGAAAGGCTGCTTC 840
QY 841 ACCCGAGGCAAGACATCGCTTGTCTCGACGTTGCTTCTTCTGAGTTCCTCAAGGAC 900
DB 841 ACCCGAGGCAAGACATCGCTTGTCTCGACGTTGCTTCTTCTGAGTTCCTCAAGGAC 900
QY 901 GGCACCTTACCACCTTCGAAGGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCT 960
DB 901 GGCACCTTACCACCTTCGAAGGTGGCCAGCACTCCGCGAGCTGAGATGGCAACGTTTACGCT 960
QY 961 GAGTCTGTTGAGGTTACCAATCGTTCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
DB 961 GAGTCTGTTGAGGTTACCAATCGTTCATCGAGGACCCACTCGAGGAAGATGACTGG 1020
QY 1021 GAGGTTTACCAACCTTCAGGCGCTTGAAGGAGGCGCATCGCTAAGAAAGGCTGCGCACTCC 1080
DB 1021 GAGGTTTACCAACCTTCAGGCGCTTGAAGGAGGCGCATCGCTAAGAAAGGCTGCGCACTCC 1080
QY 1081 TTCTTCTGTCACCAACCTTCAGGCGCTTGAAGGAGGCGCATCGCTAAGAAAGGCTGCGCACTCC 1140
DB 1081 TTCTTCTGTCACCAACCTTCAGGCGCTTGAAGGAGGCGCATCGCTAAGAAAGGCTGCGCACTCC 1140
QY 1141 ATCTGTTTAAAGTGAACAGATCGGTACCTTCACGAGACCTTCGACGCTGTCGACATG 1200
DB 1141 ATCTGTTTAAAGTGAACAGATCGGTACCTTCACGAGACCTTCGACGCTGTCGACATG 1200
QY 1201 GCTCACCGCGAGGCTTACATCTCATGATGTCACCGTTCGCTGAGACCGGAGCAC 1260
DB 1201 GCTCACCGCGAGGCTTACATCTCATGATGTCACCGTTCGCTGAGACCGGAGCAC 1260
```

```
QY 1261 ACCATTGCTGACCTCGCAGTTGACCTCAACTGTGGCAGATCAAGACTGGTGTCTCCAGCA 1320
DB 1261 ACCATTGCTGACCTCGCAGTTGACCTCAACTGTGGCAGATCAAGACTGGTGTCTCCAGCA 1320
QY 1321 CGTTCCGACCGGTTCGCAAAAGTACAACAGCTTCTCCGCATCCGAGAGCTGCTTGGCGAC 1380
DB 1321 CGTTCCGACCGGTTCGCAAAAGTACAACAGCTTCTCCGCATCCGAGAGCTGCTTGGCGAC 1380
QY 1381 GCCGGCGTCTACGAGGTTCGCGCATTCGCCAGCTTTCAGGCTAAATAAAAGCGCTT 1440
DB 1381 GCCGGCGTCTACGAGGTTCGCGCATTCGCCAGCTTTCAGGCTAAATAAAAGCGCTT 1440
QY 1441 TTCGACGCCCGGTAAACCTCAAGGTTGCGGGCGTTCGCTTACTACTGTTACTGTTGT 1500
DB 1441 TTCGACGCCCGGTAAACCTCAAGGTTGCGGGCGTTCGCTTACTACTGTTACTGTTGT 1500
QY 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTATAAAGGCTTGTTCCTGCTCT 1560
DB 1501 GACTATGATCGAGGATTTATGGCAAGCAGAAAGAACTATAAAGGCTTGTTCCTGCTCT 1560
QY 1561 CAAGCAGGGAACGCTGCTT 1578
DB 1561 CAAGCAGGGAACGCTGCTT 1578

RESULT 3
US-09-860-768-1
; Sequence 1, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-1

Query Match 99.9%; Score 1576.4; DB 9; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTGGGATATGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
DB 1 GGCTGGGATATGGTAGTTTTCGCCACTAATTTCAACTGATTCGCTCATCGAAACAAGA 60
QY 61 TTGCTGCAACAATTTGGGTGTAGACGATTTGAAGACATTTGATCAGTGAATTAATTCCTAG 120
DB 61 TTGCTGCAACAATTTGGGTGTAGACGATTTGAAGACATTTGATCAGTGAATTAATTCCTAG 120
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
DB 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTTATTCGCTCGC 180
QY 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGGTTTTCCTGGATGACGGT 240
DB 181 GAAATTTCTGACTCCCGCGGTAAACCAACCGTTCGAGGAGGTTTTCCTGGATGACGGT 240
QY 241 TCCACCGGTTCGAGGTTCCTCATCGGCGCATCCACCGGCTCCACCGAGGCTCATGAG 300
DB 241 TCCACCGGTTCGAGGTTCCTCATCGGCGCATCCACCGGCTCCACCGAGGCTCATGAG 300
QY 301 CTGCGTGAACGGTGGCGATCGCTACCTGGGCAAGGGCGTTTTGAAGGCGAGTTGAAAACGTC 360
```

```
Db 301 |||||CTGGTGACGGTGGCGATCGTACTCTGGCGAAGGGCGTTTTGAAGGAGATTGAAAAAGTCTC 360
Qy 361 AACGAAGAAATCGCGCAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
Db 361 AACGAAGAAATCGCGCAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGATGGCACCAGCCNAAGAATCCCGCTGGGTGGGTGAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGATGGCACCAGCCNAAGAATCCCGCTGGGTGGGTGAAACGCAATC 480
Qy 481 CTTGGTGTTCATGGCTGTGTCAAAGGCTGCTCTGATTCGCGAGGCTCCCACTGCTC 540
Db 481 CTTGGTGTTCATGGCTGTGTCAAAGGCTGCTCTGATTCGCGAGGCTCCCACTGCTC 540
Qy 541 CGCTACATCGGTGGACCAAAAGCGACACGTTCTTCCAGTTCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAAAAGCGACACGTTCTTCCAGTTCCAATGATGAACATCATCAAC 600
Qy 601 GGTGGCGCTCAGCGTGAAGCTCGGTGTTGACGTTTCAAGAAATTCATGATCGCTCCATCGGT 660
Db 601 GGTGGCGCTCAGCGTGAAGCTCGGTGTTGACGTTTCAAGAAATTCATGATCGCTCCATCGGT 660
Qy 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGCTTACCACGCACTGAAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGCTTACCACGCACTGAAGTCC 720
Qy 721 GTCATCAAGAAAGGCGCTGTCCACCGACTTGGCGATGAGGGCGCTTCGCTCCTTCC 780
Db 721 GTCATCAAGAAAGGCGCTGTCCACCGACTTGGCGATGAGGGCGCTTCGCTCCTTCC 780
Qy 781 GTGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
Db 781 GTGGCTCCACCGCTGAGGCTCTTGACCTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
Qy 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCAGGCAAGGACATCGCTCTTGCTCTGGACGTTGCTTCTCTGAGTTCTTCAAGGAC 900
Qy 901 GGCACCTTACCACCTCGAAGGTGGCCAGCACTCCGCACTGAGATGGCAAAAGTTAGCT 960
Db 901 GGCACCTTACCACCTCGAAGGTGGCCAGCACTCCGCACTGAGATGGCAAAAGTTAGCT 960
Qy 961 GAGCTCGTTGACGGTACCACCTGCTCCATCGAGACCCACTCGAGGAAGATGACTGG 1020
Db 961 GAGCTCGTTGACGGTACCACCTGCTCCATCGAGACCCACTCGAGGAAGATGACTGG 1020
Qy 1021 GAGGTTTACACCACTTCAACCGCAACCATCGGCGAACAAGTTTCAAGTTTGGCGACGAC 1080
Db 1021 GAGGTTTACACCACTTCAACCGCAACCATCGGCGAACAAGTTTCAAGTTTGGCGACGAC 1080
Qy 1081 TTCTTGTGTCACCAACCTCGAGCGCTGAAGAGGGCATCGCTAAGAAGGCTGCAACTCC 1140
Db 1081 TTCTTGTGTCACCAACCTCGAGCGCTGAAGAGGGCATCGCTAAGAAGGCTGCAACTCC 1140
Qy 1141 ATCTGTGTTAAGTTGAACCAAGATCGGTACCTTCCACGAGACCTTCCACGCTGTCGACATG 1200
Db 1141 ATCTGTGTTAAGTTGAACCAAGATCGGTACCTTCCACGAGACCTTCCACGCTGTCGACATG 1200
Qy 1201 GCTCACGGCGAGGCTACACTCGATGATGTCCACCGTTCCGGTGAGACCGAGGACACC 1260
Db 1201 GCTCACGGCGAGGCTACACTCGATGATGTCCACCGTTCCGGTGAGACCGAGGACACC 1260
Qy 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCCAGCA 1320
Qy 1321 CGTTCCGACCGGTTCGCAAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGGTTCGCAAAAGTACAAACAGCTTCTCCGATCGAGCAGCTGCTTGGCGAC 1380
Qy 1381 GCCGGGCTCTACCGAGTTCGCGGCAATCCCAAGCTTTCAGGGCTTAATTAAGCGCTT 1440
Db 1381 GCCGGGCTCTACCGAGTTCGCGGCAATCCCAAGCTTTCAGGGCTTAATTAAGCGCTT 1440
```

```
Db 1381 GCCGGGCTCTACCGAGTTCGCGGCAATCCCAAGCTTTCAGGGCTTAATTAAGCGCTT 1440
Qy 1441 TTCCAGCGCCCGGTAAACCTCAAGGTTTCGCGGGCTCGTTGCTTACTACTGTTACTGGTGT 1500
Db 1441 TTCCAGCGCCCGGTAAACCTCAAGGTTTCGCGGGCTCGTTGCTTACTACTGTTACTGGTGT 1500
Qy 1501 GACTATGATCGAGGATTTATGGCAAGAGCAAGAAAACTATAAAGGCTTGTTCCTGTCT 1560
Db 1501 GACTATGATCGAGGATTTATGGCAAGAGCAAGAAAACTATAAAGGCTTGTTCCTGTCT 1560
Qy 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578

RESULT 4
US-10-728-947-1
; Sequence 1, Application US/10728947
; Publication No. US20040220394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCES: 21123/278404
; CURRENT APPLICATION NUMBER: US/10/728,947
; PRIOR FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-1

Query Match 99.9%; Score 1576.4; DB 20; Length 1578;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCTGGGATATGGGTAGTATTTTCGCCACTAAATTTCAACTGATTCGCTCATCGAAACAAGA 60
Db 1 GGCTGGGATATGGGTAGTATTTTCGCCACTAAATTTCAACTGATTCGCTCATCGAAACAAGA 60
Qy 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATTCACGTGAAATTAATTTCTAG 120
Db 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTGATTCACGTGAAATTAATTTCTAG 120
Qy 121 TTAGCTTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGC 180
Db 121 TTAGCTTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGCACGTATTCGCTCGC 180
Qy 181 GAAATTCCTGACTCCCGCGTAACCCACCGTCGAGGAGAGGTTTCTCGATGACGCT 240
Db 181 GAAATTCCTGACTCCCGCGTAACCCACCGTCGAGGAGAGGTTTCTCGATGACGCT 240
Qy 241 TCCCAGGTTGTGCGAGGTTTCCATCCGCGCATCCACCGGGTCCACGAGGCTCATGAG 300
Db 241 TCCCAGGTTGTGCGAGGTTTCCATCCGCGCATCCACCGGGTCCACGAGGCTCATGAG 300
Qy 301 CTGCGTGACGGTGGCGATCGCTACTTGGGCAAGGGGCTTTTGAAGCAAGTTGAAAAAGTCTC 360
Db 301 CTGCGTGACGGTGGCGATCGCTACTTGGGCAAGGGGCTTTTGAAGCAAGTTGAAAAAGTCTC 360
Qy 361 AACGAAGAAATCGCGCAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
Db 361 AACGAAGAAATCGCGCAGAGCTCGCTGGCTAGAGCTGACGATCAGCGCTCATCGAC 420
Qy 421 GAAGCAATGATCAAGCTTGATGGCACCAGCCNAAGAATCCCGCTGGGTGGGTGAAACGCAATC 480
Db 421 GAAGCAATGATCAAGCTTGATGGCACCAGCCNAAGAATCCCGCTGGGTGGGTGAAACGCAATC 480
Qy 481 CTTGGTGTTCATGGCTGTGTCAAAGGCTGCTCTGATTCGCGAGGCTCCCACTGCTC 540
Db 481 CTTGGTGTTCATGGCTGTGTCAAAGGCTGCTCTGATTCGCGAGGCTCCCACTGCTC 540
```

Db 421 GAAAGCAATGATCAAGCTTGTATGGCACCGCCAAAGTCCGGCTGGTGCAGAAACGCAATC 480
QY 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGTGCTGATTTCCGAGGCTCCCACTGTTC 540
Db 481 CTTGGTGTTCCTATGGCTGTTCGAAAGGCTGTGCTGATTTCCGAGGCTCCCACTGTTC 540
QY 541 CGCTACATCGGTGGACCAACGCAACGCACTTCTTCCAGTTCCCAATGATGAACATCATCAAC 600
Db 541 CGCTACATCGGTGGACCAACGCAACGCACTTCTTCCAGTTCCCAATGATGAACATCATCAAC 600
QY 601 GGTGGGCTCACGCTGACTCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGT 660
Db 601 GGTGGGCTCACGCTGACTCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGT 660
QY 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCACTGAGTCC 720
Db 661 GCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGGAGGTCTACCAACGCACTGAGTCC 720
QY 721 GTCATCAAGGAAAGGCTCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCC 780
Db 721 GTCATCAAGGAAAGGCTCTGTCCACCGGACTTGGCGATGAGGGCGGCTTCGCTCTTCC 780
QY 781 GTGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGAGGCTGGCTTC 840
Db 781 GTGGCTCCACCGTGAAGGCTCTTGACCTTATCGTTAAGGCAATCGAAGAGGCTGGCTTC 840
QY 841 ACCCGAGGCAAGACATCGCTCTGTCTCGAGGTTGCTTCTCTGAGTTCTTCAAGGAC 900
Db 841 ACCCGAGGCAAGACATCGCTCTGTCTCGAGGTTGCTTCTCTGAGTTCTTCAAGGAC 900
QY 901 GGCACCTTACCCTTCAAGGTTGGCGAGCACTCCGCGAGCTGAGATGGCAACGTTACGCT 960
Db 901 GGCACCTTACCCTTCAAGGTTGGCGAGCACTCCGCGAGCTGAGATGGCAACGTTACGCT 960
QY 961 GAGCTGTTGACCGGTAACCAATCGTCTCATCGAGACCCACTGCGAGGAAGATGACTGG 1020
Db 961 GAGCTGTTGACCGGTAACCAATCGTCTCATCGAGACCCACTGCGAGGAAGATGACTGG 1020
QY 1021 GAGGTTTACACCAACTCCACCGCAACCATCGGCGACAAGGTTTCAGTCTGTGGCGAGAC 1080
Db 1021 GAGGTTTACACCAACTCCACCGCAACCATCGGCGACAAGGTTTCAGTCTGTGGCGAGAC 1080
QY 1081 TTCTTCTGTCACCAACCTGAGCGCTGAAGGAGGCACTCCTAAGAGGCTGCAACTCC 1140
Db 1081 TTCTTCTGTCACCAACCTGAGCGCTGAAGGAGGCACTCCTAAGAGGCTGCAACTCC 1140
QY 1141 ATCTGTTAAGGTGAACCAAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db 1141 ATCTGTTAAGGTGAACCAAGATCGGTACCCCTCACCGAGACCTTCGACGCTGTCGACATG 1200
QY 1201 GCTACGGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGGAGCGAGGACACC 1260
Db 1201 GCTACGGCGAGGCTACACCTCCATGATGTCCACCGTTCGGTGGAGCGAGGACACC 1260
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTGCTGCTCAGCA 1320
Db 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCGAGATCAAGACTGCTGCTCAGCA 1320
QY 1321 CGTTCCGACCGTGTCCGAAAGTACAACAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGAC 1380
Db 1321 CGTTCCGACCGTGTCCGAAAGTACAACAGCTTCTCCGCAATCGAGCAGCTGCTTGGCGAC 1380
QY 1381 GCGGCGCTTACCGAGGTCCAGGCAATCCGAGCTTTCAGGGCTTAATAAAGCGCTT 1440
Db 1381 GCGGCGCTTACCGAGGTCCAGGCAATCCGAGCTTTCAGGGCTTAATAAAGCGCTT 1440
QY 1441 TTCAGCGCCCGGTAACTCAAGGTTGCGGGCGTCTGCTTCTTACTACTGTTACTGCTGT 1500
Db 1441 TTCAGCGCCCGGTAACTCAAGGTTGCGGGCGTCTGCTTCTTACTACTGTTACTGCTGT 1500
QY 1501 GACTATGTCAGGATTTATGGCAAGCAGAGAAATCTCATAAAGGCTTGTTCCTGTCT 1560

Db 1501 GACTATGTCAGGATTTATGGCAAGCAGAGAAATCTCATAAAGGCTTGTTCCTGTCT 1560
QY 1561 CAAGCAGGGAACGTGCTT 1578
Db 1561 CAAGCAGGGAACGTGCTT 1578
RESULT 5
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Query Match 99.9%; Score 1576.4; DB 9; Length 3309400;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCGTGGGATATCGGTAGTTTTCGCCACTAAATTTCAACTGATTCGCTCATCGAACAAGA 60
Db 1034799 GCGTGGGATATCGGTAGTTTTCGCCACTAAATTTCAACTGATTCGCTCATCGAACAAGA 1034858
QY 61 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCACGCTGAATATTTCTAG 120
Db 1034859 TTCGTGCAACAATTTGGGTGTAGACGTGATTGAAGACATTTTGATCACGCTGAATATTTCTAG 1034918
QY 121 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 180
Db 1034919 TTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAAATCATGACGTATTCGCTCGC 1034978
QY 181 GAAATTCGACTCCCGCGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGACGCT 240
Db 1034979 GAAATTCGACTCCCGCGTAAACCAACCGTCGAGGACAGGTTTTCCTGGATGACGCT 1035038
QY 241 TCCACCGGTGTCCAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACAGGGCTCATGAG 300
Db 1035039 TCCACCGGTGTCCAGGTGTTCCATCCGGCGCATCCACCGCGCTCCACAGGGCTCATGAG 1035098
QY 301 CTGCGTGA CGGTGGCGATCGCTACCTTGGGCAAGGGGCTTTTGAAGGAGTTGAAACGTC 360
Db 1035099 CTGCGTGA CGGTGGCGATCGCTACCTTGGGCAAGGGGCTTTTGAAGGAGTTGAAACGTC 1035158
QY 361 AACGAGAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACCATCAGCGCTCATCGAC 420
Db 1035159 AACGAGAATCGGCGACGAGCTCGCTGGCTAGAGGCTGACCATCAGCGCTCATCGAC 1035218
QY 421 GAAGCAATGATCAAGCTTGTATGGCAACCGCCAAAGTCCCGCTGGGTGCAAAACGCAATC 480

Db 1035219 GAAGCAATGATCAAGCTTGATGGCACCGCCAAACAGTCCCGCCTGGGTGCAGAACGCATC 1035278
QY 481 CTTGGTGTTCCTATGGCTGTGTCGATGTCGAGAGGCTCCCACTGTTTC 540
Db 1035279 CTTGGTGTTCCTATGGCTGTGTCGATGTCGAGAGGCTCCCACTGTTTC 1035338
QY 541 CGCTACATCGGTGACCAACGACACAGTTCCTCCAGTTCCCAATGATGAACATCATCAAC 600
Db 1035339 CGCTACATCGGTGACCAACGACACAGTTCCTCCAGTTCCCAATGATGAACATCATCAAC 1035398
QY 601 GGTGGCGCTCACGCTGACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGT 660
Db 1035399 GGTGGCGCTCACGCTGACTCCGGTGTGAGCTTCAGGAATTCATGATCGCTCCAATCGGT 1035458
QY 661 GCAGAGACCTTCCTGAGGCTCTCCGCAAGGGCGAGGCTTACACGCACTGAAGTCC 720
Db 1035459 GCAGAGACCTTCCTGAGGCTCTCCGCAAGGGCGAGGCTTACACGCACTGAAGTCC 1035518
QY 721 GTCATCAAGAAAAGGCGCTGTCCACCGACTTGGCGATCAGGCGGCTTCGCTCCTTCC 780
Db 1035519 GTCATCAAGAAAAGGCGCTGTCCACCGACTTGGCGATCAGGCGGCTTCGCTCCTTCC 1035578
QY 781 GTGGCTCCACCGCTGAGGCTCTTGACTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 840
Db 1035579 GTGGCTCCACCGCTGAGGCTCTTGACTTATCGTTAAGCAATCGAAGAGGCTGGCTTC 1035638
QY 841 ACCCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTGAGTTCCTCAAGGAC 900
Db 1035639 ACCCCAGGCAAGGACATCGCTCTTGCTCTGAGAGTTCCTGAGTTCCTCAAGGAC 1035698
QY 901 GGCACCTTACCACCTTCGAAGGTGGCGACACTCCGCACTGAGATGGCAACGTTTACGCT 960
Db 1035699 GGCACCTTACCACCTTCGAAGGTGGCGACACTCCGCACTGAGATGGCAACGTTTACGCT 1035758
QY 961 GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGAACCCACTGAGGAAGATGACTGG 1020
Db 1035759 GAGCTCGTTGACCGGTACCCAAATCGTCTCCATCGAGAACCCACTGAGGAAGATGACTGG 1035818
QY 1021 GAGGTTTACACCAACCTCACCAGCAACCATCGGCAAGGTTTACGATCGTTGGCGGAC 1080
Db 1035819 GAGGTTTACACCAACCTCACCAGCAACCATCGGCAAGGTTTACGATCGTTGGCGGAC 1035878
QY 1081 TTCTTCGTACCAACCTCGAGCGCTGAAGAGGGCATCGCTAAGAGGCTGCAACTCC 1140
Db 1035879 TTCTTCGTACCAACCTCGAGCGCTGAAGAGGGCATCGCTAAGAGGCTGCAACTCC 1035938
QY 1141 ATCCTGGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1200
Db 1035939 ATCCTGGTTAAGGTGAACCAAGATCGGTACCTCACCGAGACCTTCGACGCTGTCGACATG 1035998
QY 1201 GCTCACGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCTGAGACCGAGGACCC 1260
Db 1035999 GCTCACGCGCAGGCTACACCTCCATGATGTCCACCGTTCGCTGAGACCGAGGACCC 1036058
QY 1261 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTCTCCAGCA 1320
Db 1036059 ACCATTGCTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGTGTCTCCAGCA 1036118
QY 1321 CGTTCCGACCGTGTGCGAAGTACACACGACTTCTCGGATCGAGCAGCTGCTTGGCGAC 1380
Db 1036119 CGTTCCGACCGTGTGCGAAGTACACACGACTTCTCGGATCGAGCAGCTGCTTGGCGAC 1036178
QY 1381 GCCGGCGTCTACGAGGTGCGACGATTCCTCCACGCTTTCAGGGCTAAATAAAGCGCTT 1440
Db 1036179 GCCGGCGTCTACGAGGTGCGACGATTCCTCCACGCTTTCAGGGCTAAATAAAGCGCTT 1036238
QY 1441 TTCGACCGCGGTAACTCAAGGTTGCGCGGCGTGTGCTTACTTACTGTTACTGTTGT 1500
Db 1036239 TTCGACCGCGGTAACTCAAGGTTGCGCGGCGTGTGCTTACTTACTGTTACTGTTGT 1036298
QY 1501 GACTATGATCGAGGATATGGCAAGCAAGCAAGAACTCATAAAGGCTGTGCTCTGTCT 1560

Db 1036299 GACTATGATCGAGGATATGGCAAGCAAGAACTCATAAAGGCTGTGCTCTGTCT 1036358
QY 1561 CAAGCAGGGAACGTGCTT 1578
Db 1036359 CAAGCAGGGAACGTGCTT 1036376
RESULT 6
US-10-494-836-7
; Sequence 7, Application US/10494836
; Publication No. US20050014233A1
; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Hartwig
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopprogge, Corinna
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy product
; FILE REFERENCE: BGI-167US
; CURRENT APPLICATION NUMBER: US/10/494,836
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: PCT/EP02/12135
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 101 54 270.4
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 7
; LENGTH: 1405
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1375)
; OTHER INFORMATION: RXA00235
US-10-494-836-7

Query Match 88.8%; Score 1401.8; DB 21; Length 1405;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 51 CGAAACAAGATTCTGTGTCGCAACAATTGGGTGTAGACGCTGATTGAAGACATTTGATCAGCTGA 110
Db 1 CGAAACAAGATTCTGTGTCGCAACAATTGGGTGTAGACGCTGATTGAAGACATTTGATCAGCTGA 60
QY 111 ATAAATCTAGTTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCACGT 170
Db 61 ATAAATCTAGTTAGCTCCCAAGTTGGCATAGAGGCCACAGTGGCTGAAATCATGCACGT 120
QY 171 ATTTCGCTCGCGAAATTTCTGCAGCTCCCGCGTAAACCCACCGTCGAGGCAGAGGTTTTCCT 230
Db 121 ATTTCGCTCGCGAAATTTCTGCAGCTCCCGCGTAAACCCACCGTCGAGGCAGAGGTTTTCCT 180
QY 231 GGATGACGTTCCCAAGTGTGCGAGGTTTCCATCCCGCGCATCCACCGGGCTCCACGA 290
Db 181 GGATGACGTTCCCAAGTGTGCGAGGTTTCCATCCCGCGCATCCACCGGGCTCCACGA 240
QY 291 GGCTCATGAGCTGCGTGAGCGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGT 350
Db 241 GGCTCATGAGCTGCGTGAGCGTGGCGATCGCTACCTGGGCAAGGGCGTTTGAAGGCAGT 300
QY 351 TGAAACAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACGATCAGCG 410
Db 301 TGAAACAAGTCAACGAAGAAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACGATCAGCG 360
QY 411 CTTCAATCGAAGCAATGATCAAGTTCGATGGCAACCGCCAAACAAGTCCCGCTGGGTGC 470
Db 361 CTTCAATCGAAGCAATGATCAAGTTCGATGGCAACCGCCAAACAAGTCCCGCTGGGTGC 420
QY 471 AAACCCAAATCCTTGGTGTTCATCGCTGTTGCAAGGCTGCTGATTCGACAGGCT 530
Db 421 AAACCCAAATCCTTGGTGTTCATCGCTGTTGCAAGGCTGCTGATTCGACAGGCT 480

```
QY 531 CCCACTGTTCCGCTACATCGGTGGACCAAAACGACACAGTTCCTTCCAGTTCCTCAATGATGAA 590
Db 481 CCCACTGTTCCGCTACATCGGTGGACCAAAACGACACAGTTCCTTCCAGTTCCTCAATGATGAA 540
QY 591 CATCATCAAGCGTGGCGCTCAGCTGACTCCGGTGTGAGTTCAGGAATTCATGATCGC 650
Db 541 CATCATCAAGCGTGGCGCTCAGCTGACTCCGGTGTGAGTTCAGGAATTCATGATCGC 600
QY 651 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACCAACGC 710
Db 601 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGGAGGTCTACCAACGC 660
QY 711 ACTGAAGTCCGTATCAAGAAAGAGCGCTGTCCACCGGACTTTGGCGATGAGGCGCGCTT 770
Db 661 ACTGAAGTCCGTATCAAGAAAGAGCGCTGTCCACCGGACTTTGGCGATGAGGCGCGCTT 720
QY 771 CGCTCCCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
Db 721 CGCTCCCTTCGCTCGGCTCCACCGGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780
QY 831 GGCTGGCTTCACCCCGGCAAGGACATCGCTCTTGCTCTGAGCTTGCTTCTCTGAGTT 890
Db 781 GGCTGGCTTCACCCCGGCAAGGACATCGCTCTTGCTCTGAGCTTGCTTCTCTGAGTT 840
QY 891 CTTCAAGGACGGACCTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAA 950
Db 841 CTTCAAGGACGGACCTTACCACTTCGAAGGTGGCCAGCACTCCGACGCTGAGATGGCAAA 900
QY 951 CGTTTAGCTGAGCTGTTGACGGGTACCAATCGTCTCATCGAGGACCCACTGCAGGA 1010
Db 901 CGTTTAGCTGAGCTGTTGACGGGTACCAATCGTCTCATCGAGGACCCACTGCAGGA 960
QY 1011 AGATGACTGGAGGGTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTCCAGATCGT 1070
Db 961 AGATGACTGGAGGGTTACACCAACCTCACCGCAACCATCGGCGCAAGGTTCCAGATCGT 1020
QY 1071 TGGCGAGCACTTCTTGCTACCAACCTTGAGCGCTGAAGGAGGATCGCTAAGAAGGC 1130
Db 1021 TGGCGAGCACTTCTTGCTACCAACCTTGAGCGCTGAAGGAGGATCGCTAAGAAGGC 1080
QY 1131 TGGCAACTCATCTGTTAAGGTGAACAGATCGGTACCTTCACCGAGACCTTTCGACGC 1190
Db 1081 TGGCAACTCATCTGTTAAGGTGAACAGATCGGTACCTTCACCGAGACCTTTCGACGC 1140
QY 1191 TGTGCAATGGCTCACCGCGAGGCTACACCTCCATGATGTCCTCCAGCTTCGCGTGAGAC 1250
Db 1141 TGTGCAATGGCTCACCGCGAGGCTACACCTCCATGATGTCCTCCAGCTTCGCGTGAGAC 1200
QY 1251 CGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1310
Db 1201 CGAGGACACCACTTGTGACCTCGCAGTTGCACTCAACTGTGGCCAGATCAAGACTGG 1260
QY 1311 TGCTCCAGCAGCTTCGACCGGTGCGAAAGTCAACAGCTTCTCGCATTCGAGCAGCT 1370
Db 1261 TGCTCCAGCAGCTTCGACCGGTGCGAAAGTCAACAGCTTCTCGCATTCGAGCAGCT 1320
QY 1371 GCTTGGCGAGCGCGGCTTACCGAGGTCCAGGGATTCCTCCAGCTTTCAGGGCTAAAT 1430
Db 1321 GCTTGGCGAGCGCGGCTTACCGAGGTCCAGGGATTCCTCCAGCTTTCAGGGCTAAAT 1380
QY 1431 AAAAGCGCTTTTCGACGCCCGGTAA 1455
Db 1381 AAAAGCGCTTTTCGACGCCCGGTAA 1405
```

RESULT 7

```
US-10-781-014-71
; Sequence 71, Application US/10781014
; Publication No: US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
```

```
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: BGI-126CPCN
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 05/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/151,572
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19931412.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931424.1
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931428.4
; PRIOR FILING DATE: 1999-07-08
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 784
; SEQ ID NO 71
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1375)
; OTHER INFORMATION: RXA00235
; US-10-781-014-71
```

```
Query Match 88.4%; Score 1394.8; DB 19; Length 1398;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1396; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 CGAAACAAGATTCGTCAACAATTGGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 110
Db 1 CGAAACAAGATTCGTCAACAATTGGGTGTAGACGTGATTGAAGACATTTGATCACGTGA 60

QY 111 ATAATTCTAGTTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGT 170
Db 61 ATAATTCTAGTTAGCTCCCAAGTTGGCATAGGAGGCCACAGTGGCTGAATCATGCACGT 120

QY 171 ATTGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCAACCGTTCGAGCAGAGGTTTCT 230
Db 121 ATTGCTCGCGAAATTTCTCGACTCCCGCGGTAAACCAACCGTTCGAGCAGAGGTTTCT 180

QY 231 GGATGAGCGTTCACCGGTGTCCAGGTGTTCATCGGCGCATCCACCGGCTCCACGA 290
Db 181 GGATGAGCGTTCACCGGTGTTCATCGGCGCATCCACCGGCTCCACGA 240

QY 291 GGCTCATGAGTTCGCTGACGTCGGCGATCGTACCTGGGCAAGGGCTTTTGAAGGCGT 350
Db 241 GGCTCATGAGTTCGCTGACGTCGGCGATCGTACCTGGGCAAGGGCTTTTGAAGGCGT 300

QY 351 TGAACAACGTCAACGAAGAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACGATCAGCG 410
Db 301 TGAACAACGTCAACGAAGAATCGGCGACGAGCTCGCTGGCCTAGAGGCTGACGATCAGCG 360

QY 411 CCTCATCGAGGAAGCAATGATCAAGCTTTCAGCTTTCGACCGGCAACCAAGTCCGCGCTGGGTGC 470
Db 361 CCTCATCGAGGAAGCAATGATCAAGCTTTCAGCTTTCGACCGGCAACCAAGTCCGCGCTGGGTGC 420

QY 471 AAACGCAATCTTGGTGTTCATGCTGTTGCAAGGCTGCTGCTGATTCCGCGAGGCT 530
```



```

Db 421 AAAGCAATCTTGGTGTTCATGCTGTGCAAGGCTGCTGCTGATTCGAGGCT 480
QY 531 CCACATGTTCCGCTACATCGGTGGAACAAACGACAGCTTCTCCAGTTCGAATGATGAA 590
Db 481 CCACATGTTCCGCTACATCGGTGGAACAAACGACAGCTTCTCCAGTTCGAATGATGAA 540
QY 591 CATCATCAACGGTGGGCTCACCTGACTCGGTGTTGAGGTTTCAGGAATTCATGATCGC 650
Db 541 CATCATCAACGGTGGGCTCACCTGACTCGGTGTTGAGGTTTCAGGAATTCATGATCGC 600
QY 651 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGC 710
Db 601 TCCAATCGGTGCAGAGACCTTCTCTGAGGCTCTCCGCAACGGCGCGAGGTCTACCAACGC 660
QY 711 ACTGAAGTCCGTTCATCAGGAAGAGGCTGTGCAACGGGACTTGGCGATGAGGGCGGCTT 770
Db 661 ACTGAAGTCCGTTCATCAGGAAGAGGCTGTGCAACGGGACTTGGCGATGAGGGCGGCTT 720
QY 771 CGCTCCCTTCGCTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 830
Db 721 CGCTCCCTTCGCTCGGCTCCACCGTGAGGCTCTTGACCTTATCGTTAAGGCAATCGAGAA 780
QY 831 GGCTGGCTTCACCCGAGGCAAGGACATCGCTCTTGCTCTGGAGCTTGTCTTCTCTGAGTT 890
Db 781 GGCTGGCTTCACCCGAGGCAAGGACATCGCTCTTGCTCTGGAGCTTGTCTTCTCTGAGTT 840
QY 891 CTTCAAGGACGGACCTTACACATTCGAAAGTGGCGAGCACTCCGCGAGCTGAGATGGCAAA 950
Db 841 CTTCAAGGACGGACCTTACACATTCGAAAGTGGCGAGCACTCCGCGAGCTGAGATGGCAAA 900
QY 951 CGTTTACGCTGAGCTGTTGACCGGTACCCCAATCGCTCTCCATCGAGGAGGCGATCGCT 1010
Db 901 CGTTTACGCTGAGCTGTTGACCGGTACCCCAATCGCTCTCCATCGAGGAGGCGATCGCT 960
QY 1011 AGATGACTGGAGGGTTACACCAACCTCACCGCAACCATCGGCGCAAGGGTTCCAGATCGT 1070
Db 961 AGATGACTGGAGGGTTACACCAACCTCACCGCAACCATCGGCGCAAGGGTTCCAGATCGT 1020
QY 1071 TGGCGAGCACTTCTGTCACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAGGC 1130
Db 1021 TGGCGAGCACTTCTGTCACCAACCTGAGCGCTGAGGAGGCGATCGCTAAGAGGC 1080
QY 1131 TGCCAACTCCATCTGTTAAGGTGACACAGATCGGTACCTCCACGAGACCTTCCAGCGC 1190
Db 1081 TGCCAACTCCATCTGTTAAGGTGACACAGATCGGTACCTCCACGAGACCTTCCAGCGC 1140
QY 1191 TGTGCAATGCTCACCGCGAGGCTACACCTCCATGATGTCACCGTTCGCGTGAGAC 1250
Db 1141 TGTGCAATGCTCACCGCGAGGCTACACCTCCATGATGTCACCGTTCGCGTGAGAC 1200
QY 1251 CGAGGACACCACTTGTGACCTCGAGTTCGACTTCACTGTCGAGTCAAGACTGG 1310
Db 1201 CGAGGACACCACTTGTGACCTCGAGTTCGACTTCACTGTCGAGTCAAGACTGG 1260
QY 1311 TGCTCCAGCACTTCCGACCGTTCGCAAGTACACCACTTCTCCGATTCAGCAGCT 1370
Db 1261 TGCTCCAGCACTTCCGACCGTTCGCAAGTACACCACTTCTCCGATTCAGCAGCT 1320
QY 1371 GCTTGGCGAGCGCGGCTTACCGAGGTGCGAGGATTCGCCAGCTTTCAGGGCTTAAAT 1430
Db 1321 GCTTGGCGAGCGCGGCTTACCGAGGTGCGAGGATTCGCCAGCTTTCAGGGCTTAAAT 1380
QY 1431 AAAAGCGCTTTCGACGC 1448
Db 1381 AAAAGCGCTTTCGACGC 1398

```

RESULT 8
 US-09-738-626-1085
 ; Sequence 1085, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1085
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1085

Query Match      80.7%; Score 1273.4; DB 9; Length 1275;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 151 GTGGCTGAAATCATGACAGTATTCGCTCCGAAAATTCGACTCCCGCGGTAAACCAACC 210
Db 1 GTGGCTGAAATCATGACAGTATTCGCTCCGAAAATTCGACTCCCGCGGTAAACCAACC 60
QY 211 GTCGAGGACAGAGGTTTTCTGGATGACGGTTCACCGGTGTCGACGGTGTTCATCCGGC 270
Db 61 GTCGAGGACAGAGGTTTTCTGGATGACGGTTCACCGGTGTCGACGGTGTTCATCCGGC 120
QY 271 GCATCCACCGGGTCCACGAGGCTCATGAGCTGCGTGAAGGATCGCTACCTGGGC 330
Db 121 GCATCCACCGGGTCCACGAGGCTCATGAGCTGCGTGAAGGATCGCTACCTGGGC 180
QY 331 AAGGGCGTTTTGAAGGAGTTGAACAGTCAACGAGAAATCGGCGACGAGCTCGGTGGC 390
Db 181 AAGGGCGTTTTGAAGGAGTTGAACAGTCAACGAGAAATCGGCGACGAGCTCGGTGGC 240
QY 391 CTAGAGGCTGACGATCAGGCGCTCATCGACGAGCAATGATCAAGCTTGTATGGCACCGCC 450
Db 241 CTAGAGGCTGACGATCAGGCGCTCATCGACGAGCAATGATCAAGCTTGTATGGCACCGCC 300
QY 451 ACAAAGTCCGGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTCGAAAGGCT 510
Db 301 ACAAAGTCCGGCTGGGTGCAACGCAATCTTGGTGTTCATGCTGTTCGAAAGGCT 360
QY 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACAGCTT 570
Db 361 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAACGACAGCTT 420
QY 571 CTTTCCAGTTCCCAATGATGAACATCATCAACGGTGGGCTCACGCTGACCTCCGGTGTGAC 630
Db 421 CTTTCCAGTTCCCAATGATGAACATCATCAACGGTGGGCTCACGCTGACCTCCGGTGTGAC 480
QY 631 GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGGAATTCATGATCGCTCCAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 540
QY 691 GCGCGGAGGCTTACCAACGACCTGAAGTCCGTCATCAAGGAAAGGGCTGTCCACCGGA 750
Db 541 GCGCGGAGGCTTACCAACGACCTGAAGTCCGTCATCAAGGAAAGGGCTGTCCACCGGA 600
QY 751 CTTGGCGATGAGGGGCGGCTTTCGCTCCTTCGCTCGGCTCCACCGTGAGGCTCTTGACCTT 810

```

```
Db 601 CTTGGCATGAGGCGGCTTCGCTCCCTTCGCTGGCTCCACCGTGAGGCTCTTGACCTT 660
Qy 811 ATCTTTAAGGCAATCGAGAAGGCTGGCTTACCCCGAGGACATCGCTCTTGCTCTG 870
Db 661 ATCTTTAAGGCAATCGAGAAGGCTGGCTTACCCCGAGGACATCGCTCTTGCTCTG 720
Qy 871 GACCTGCTCTCTGAGTCTTCAAGGACGGCACCTACCACTTCGAAGGTGGCCAGCAC 930
Db 721 GACCTGCTCTCTGAGTCTTCAAGGACGGCACCTACCACTTCGAAGGTGGCCAGCAC 780
Qy 931 TCCGACGCTGAGTGGCAAGGCTTACGCTGAGCTCGTTGACGGGTACCCCAATCGTCTCC 990
Db 781 TCCGACGCTGAGTGGCAAGGCTTACGCTGAGCTCGTTGACGGGTACCCCAATCGTCTCC 840
Qy 991 ATCGAGGACCCATCGCAGGAGATGACTGGGAGGTTACACCAACCTCACCGCAACCATC 1050
Db 841 ATCGAGGACCCATCGCAGGAGATGACTGGGAGGTTACACCAACCTCACCGCAACCATC 900
Qy 1051 GCGCAAGGTTGAGTGGCGAGCACTTCTTCGTACCAACCTCGAGCGCTGAAG 1110
Db 901 GCGCAAGGTTGAGTGGCGAGCACTTCTTCGTACCAACCTCGAGCGCTGAAG 960
Qy 1111 GAGGCACTCGTAGAAGGCTGCCAATCCATCTCGTTAAGTTGAACGAGATCGGTACC 1170
Db 961 GAGGCACTCGTAGAAGGCTGCCAATCCATCTCGTTAAGTTGAACGAGATCGGTACC 1020
Qy 1171 CTCACCGAGACCTTCGAGCGCTGTGCACATGCTCACCGCGAGGCTACACCTCCATGATG 1230
Db 1021 CTCACCGAGACCTTCGAGCGCTGTGCACATGCTCACCGCGAGGCTACACCTCCATGATG 1080
Qy 1231 TCCACCGCTTCGGTGAGACCGAGGACACCAACCTTCTGACCTCGAGTTGCACTCAAC 1290
Db 1081 TCCACCGCTTCGGTGAGACCGAGGACACCAACCTTCTGACCTCGAGTTGCACTCAAC 1140
Qy 1291 TGTGGCAGATCAGACTGTGTCTCCAGCAGTTCCAGCGTGTCCGAAAGTACAAACGAG 1350
Db 1141 TGTGGCAGATCAGACTGTGTCTCCAGCAGTTCCAGCGTGTCCGAAAGTACAAACGAG 1200
Qy 1351 CTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTGTACGCGAGTTCGCGAGGCATTC 1410
Db 1201 CTTCTCCGATCGAGCAGCTGCTTGGCGAGCGCGGCTGTACGCGAGTTCGCGAGGCATTC 1260
Qy 1411 CCACGCTTCAGGCG 1425
Db 1261 CCACGCTTCAGGCG 1275
```

RESULT 9

```
US-10-282-122A-17724
; Sequence 17724, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Yamamoto, Robert
; APPLICANT: Carr, Grant
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17724
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17724
```

```
Query Match 57.5%; Score 907; DB 17; Length 1275;
Best Local Similarity 82.0%; Pred. No. 1.9e-268;
Matches 1045; Conservative 0; Mismatches 230; Indels 0; Gaps 0;
Qy 151 GTGGCTGAATCATGACGATTTCCGTCGCGAAATTCGACATCCCGGGTAACCCAAAC 210
Db 1 GTGGCTGACATTTATGACGATTTTGTCTGCTGTAATCTTGACATCCCGGGTAACCCAAAC 60
Qy 211 GTGAGGACGAGGTTTCTTGGATGACGTTCCACGCTGTCGAGGTGTCGAGGTGTTCCATCCGGC 270
Db 61 GTTGAAGCCGAAGTTTCTTGGATGACGATCCACGCGTTGCGGGGTTCTTCCGGT 120
Qy 271 GCATCCACCGCGCTCCACGAGGCTCATGAGCTGCGTGACGCTGGCGATCGCTACTCGGC 330
Db 121 GCCTCCACCGGTGTTACGAGGCTCAGAGCTTCGCGAGCGTGGCGAGCGCTACTCGGC 180
Qy 331 AAGGCGTTTTTGAAGGCGTTGAAACGTCACGAAAGTTCGCGAGTTCGCGAGTTCGCTGGC 390
Db 181 AAGGCGTTCTCAATGCACTTAACAGCTGAACGAGGAAATCGCTGACGCAATCGCTGGT 240
Qy 391 CTAGAGGCTGACGATCAGCGCTCATCGACGAGCAATGATCAAGTTGATGGCACCGCC 450
Db 241 GCAGAAGCCGACGATCAGCGCTGATCGATCAGCGGATGATGCTCTCGACGGCACTGAG 300
Qy 451 AACAAAGTCCCGCTGGGTGCAAAACGCAATCCTTTGGTGTTCATGGCTGTTCGAAAGCT 510
Db 301 AACAAAGTCTGCTCGCGCAACGCTATCTTGGTGTATCCATCGCGTAGCTAAGGCT 360
Qy 511 GCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACAAACGCAACGCTT 570
Db 361 GCGCGAGAGTCTGCTGGCTTACCTTTGTACCGCTACATCGCGGCGCTTAAACGCTCAGCTT 420
Qy 571 CTTCCAGTTCCAAATGATGACATCAACGCTGGCGCTCACGCTGACTCGGTGTTGAC 630
Db 421 CTTCCAGTTCCATGATGAAATTTAAGGTTGAGGCGCACACGCTGACTCCGGGTTGAT 480
Qy 631 GTTCAGAAATTCATGATCGCTCCAAATCGGTGACAGACCTTCTCTGAGGCTCTCCGCAAC 690
Db 481 GTTCAGAGTTTCATGATGCTCTATCGGTGCCAGTCTTCTCTGAGGCTCTCGCGATG 540
Qy 691 GCGCGAGAGTCTTACCAAGCTGAAATCGGTTCATCAAGGAAAGGCGCTGTCCACCGGA 750
Db 541 GGTGACAGAGTCTACCACTCTTGAAGTCCGTTAAGTCCAAAGGAGACTTTCACCGGC 600
Qy 751 CTTGGCGATGAGGCGGCTTCCGCTTCTTCCGCTGGCTCCACCGTGGAGGCTCTTGACCTT 810
Db 601 CTCGGCGAGGAGGTTGTTTCGACCTTCTGTTGAGTCCACCAAGGAGCAGCTCTCGACCTC 660
```


Db 4372589 CGAGCGCCTGGCCCGCGGCATCGAGAGGGCTCCGCCAACGCCCTGCTCGTCAAGGTCA 4372530
QY 1157 ACCAGATCGGTACCTCACCAGACCTTTCGACGCTGTGCGACATGGGTCAACCGCGCAGGCT 1216
Db 4372529 ACCAGATCGGTTCGCTGACCGACCTTGAACCGCGTGGAGCTGGCCAGCGCAACGGCT 4372470
QY 1217 ACACCTCCATGATGTCCACCGTTCCGGTGGAGACCGAGGACACACCATTTGCTGACCTCG 1276
Db 4372469 TCAAGTGCATGATGTCCACCGCTCCGGCGAGACCGAGGACGTCAACCATCGCGACCTCG 4372410
QY 1277 CAGTTGCACCTCAACTGTGGCCAGATCAAGACTGGTCTCCAGACAGTTCCGACCGGTGCG 1336
Db 4372409 CCGTCCCGTGAATCGCGGTGAGTCAAGACCGCGCCCGCTCGGACCGGTGCG 4372350
QY 1337 CAAAGTACAAACAGCTTTCGCGATCGAGCAGCTGTGTCGCGACGCGCGGTCTACGCGAG 1396
Db 4372349 CCAAGTACAAACAGCTGCTGGGATCGAGGAGATCTTCGACGACGCGCGAGTACGCCG 4372290
QY 1397 GTCCGAGCGCATTCACACGCTTTC 1420
Db 4372289 GCCGCTCGCGCTTCCCGCGGTTC 4372266

RESULT 11

US-10-156-761-3520
; Sequence 3520, Application US/10156761
; Publication No. US20030119018A1

GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3520
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1284)
US-10-156-761-3520

Query Match

Best Local Similarity 44.1%; Score 695.4; DB 15; Length 1284;

Matches 919; Conservativity 0; Mismatches 351; Indels 3; Gaps 1;

QY 151 GTGGCTGAATCATGACGATTCGCTCGGGAATTCGACTCCCGCGGTAAACCAACC 210
Db 1 GTCCCGTCCATCAGCTGCTGAGCCCGGGAATCTCGACTCCCGAGGCAACCCGACG 60
QY 211 GTCCAGGCGAGGTTTCTCGGATGACGGTTCCACGGTCTCGAGGTGTTCATCCGGC 270
Db 61 GTCCAGGTCGAGGTTCGGCTTCGACGACGCGAGCAGCGGTCTGCGCCGCTCCGTCGGC 120
QY 271 GCATCCACCGCGTCCACGAGGCTCATAGCTCGGTGACGGTG---CGGATCGCTACCTG 327
Db 121 GCCTCCACCGGTGCCTTCGAGGGCCATCGAGCTCCGGACGGTGAACCCCAACCGTTACCAG 180
QY 328 GGCAGAGGGCTTTTGAAGGAGTTGAAAGCTCAACGAAGAAATCGCGACGAGCTCGCT 387
Db 181 GGCAAGGGGTGTCGAGAAGGCCGCTCTCGCGCTCATCGAGCAGATCGGCGCGGAGCTCGTC 240

QY 388 GGCCTAGAGGCTGACGATCAGCGCCTCATCGACGAAGCAATGATCAAGCTTGTATGGCACC 447
Db 241 GGTCTGACGCGCACCGCAGCAGCGCTGATCGACGAGCGATTTGACCTTGGACGCGCAC 300
QY 448 GCCAACAAAGTCCCGCTGGGTGCAAAAGCAATCTTGTGTGTTTCCATGGCTGTGCAAG 507
Db 301 GACAAACAAAGGCTCGCTCGGCGCAACGCCATCTCGGGGTCTCTCCCTCGCGCTCGCGCAC 360
QY 508 GGTGCTGCTGATTCGCGAGGCTTCCCACTGTTCCGCTACATCGGTGGACCAAAACGACAC 567
Db 361 GCGGCTTCGAGGCGAGCGACCTCCCGCTTTCGCTACTCTGGGCGCGCGAGCGCGCAC 420
QY 568 GTTCTTCCAGTTCCAAATGATGAACATCATCAACGCTGGCGCTCACGCTGATCTCGGTGT 627
Db 421 CTGCTGCCGCTTCGATGATGAACATCTCAACGCGGCTCGCACGCGGACTCCAACGTG 480
QY 628 GAGCTTCAGGAATTCATGATCGCTCCAAATCGGTGAGAGACCTTCTCTGAGGCTCTCCGC 687
Db 481 GACATCCAGGAGTTTATGATCGCCCGATCGGCGCGGAGTCTTCTCCGAGGCGCTCGCG 540
QY 688 AACGCGCGGAGGTCTACCAAGCACTGAAGTCCGTCTCAAGGAAAGAGGCTGTCCACC 747
Db 541 TGGGCGCGCGAGGTCTACCAACCTCAAGAGGTGCTGAGACCAAGAGGCTGTCCACC 600
QY 748 GGAATTTGGCGATGAGGCGGCTTTCGCTTCCGTCCGCTCCACCCTGAGGCTCTTTGAC 807
Db 601 GGCCTCGCGGACGAGGCGGCTTTCGCGCCCGAACCTGGAGTTCGAACCGCGCGGCTCGAC 660
QY 808 CTTATCGTTAAGCAATCGAGAAGGCTGGCTTACCCCGAGGCAAGGACATCGCTTGTCT 867
Db 661 CTCATCATCGAGGCGCATCAAGCAGGCGGTTTACATCCCGGCGAGCAGATTCGCGCTCGCG 720
QY 868 CTGGAGCTTGTCTTCTCTGAGTTCTTAAAGGACGCGACCTTACCCTTTCGAAGGTGGCCAG 927
Db 721 CTCGAGCTCGCGGCTCCGAGTCTTCAAGGACGCGAAGTACAGTTCGAGGGCAAGTCC 780
QY 928 CACTCCGAGCTGAGATGGCAACGTTTACGCTGAGCTCGTTGACCGGTACCCAAATCGTC 987
Db 781 CGCTCGGCGCGCGAGATGACCGAGTACTACGAGGAGCTCGTCTCCGCGTACCCGCTCGTC 840
QY 988 TCCATCGAGGACCCACTCGAGGAGATGACTGGGAGGTTTACACCAACCTCAGCGCAACC 1047
Db 841 TCCATCGAGGACCCCGCTGTACGAGGACGACTGGGCGCGGCTGGAGGTTCATCAGCGAAC 900
QY 1048 ATCGGCGCAAGGTTCCAGATCGTTGGCGACGACTTCTTGTTCACCAACCTTGAGCGGCTG 1107
Db 901 CTGGGCGCAAGGTCCAGATCGTGGCGGAGCCTCTTGTTCACCAACCGCGAGCGGCTG 960
QY 1108 AAGGAGGCGATCGCTAAGAGGCTGCCAACTCCATCTCGTTAAGGTGAACCAAGATCGGT 1167
Db 961 GCCCGCGCATCGAGGAGGCTCCGCCAACGCCCTGCTCGTCAAGGTCAACAGATCGGT 1020
QY 1168 ACCTCAGCGAGACTTTCGACGCTGTCGATGGCTCACCGCGCAGGCTTACCTCATG 1227
Db 1021 TCGCTCAGCGAGACCCCTGGACGCGGTTCGAGCTGGCGCGAGCGCAACGGCTTCAAGTGCATG 1080
QY 1228 ATGTCCACCGGTTCCGCTGAGCCGAGGACACCAACATTCGCTGACCTTCGAGTTGCACTC 1287
Db 1081 ATGTCCACCGCTTCCGCGAGACCGGAGGAGTCAACATCGCGACCTTCGCGCTCGCGGTG 1140
QY 1288 AACTGTGGCGCAGATCAAGACTGCTGTCGAGCAAGTTTCGACCGGTGTCGCAAAAGTACAAC 1347
Db 1141 AACTGCGGTGAGATCAAGACCGCGCGCCCGCTCGGACCGCTGTCGCGCAAGTACAAC 1200
QY 1348 CAGCTTCTCCGATTCAGAGCAGTGTGTTGGGAGCGCGGCTGTACGAGGTTCGCGAGGCA 1407
Db 1201 CAGCTCTGCGATTCGAGGAGATCTTCGACGACGCGCGGAGTACGCGCGGCTCGCGG 1260
QY 1408 TTCCCAAGCTTTC 1420
Db 1261 TTCCCGCGGTTC 1273

RESULT 12		US-10-369-493-32027	
		; Sequence 32027, Application US/10369493	
		; Publication No. US20030233675A1	
		; GENERAL INFORMATION:	
		; APPLICANT: Cao, Yongwei	
		; APPLICANT: Hinkle, Gregory J.	
		; APPLICANT: Slater, Steven C.	
		; APPLICANT: Goldman, Barry S.	
		; APPLICANT: Chen, Xianfeng	
		; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF	
		; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES	
		; FILE REFERENCE: 38-10(52052)B	
		; CURRENT APPLICATION NUMBER: US/10/369,493	
		; CURRENT FILING DATE: 2003-02-28	
		; PRIOR APPLICATION NUMBER: US 60/360,039	
		; PRIOR FILING DATE: 2002-02-21	
		; NUMBER OF SEQ ID NOS: 47374	
		; SEQ ID NO 32027	
		; LENGTH: 1269	
		; TYPE: DNA	
		; ORGANISM: Thermobifida fusca	
		US-10-369-493-32027	
		Query Match 43.0%; Score 678.2; DB 17; Length 1269;	
		Best Local Similarity 71.0%; Pred. No. 6.3e-198;	
		Matches 899; Conservative 0; Mismatches 368; Indels 0; Gaps 0;	
QY	152	TGGCTGAATCATGCGAGTATTTCGCTCGCGAAATTCGCACTCCCGGGTAACCCACCG	211
DB	2	TGGCGTTCGATCGAGGCGAGTACACGACCGGAGATTTCGACTCCCGGGTAACCCGAGG	61
QY	212	TCGAGGCGAGAGTTTTCCTGGATGACGGTTCACCGGTTCGCGAGTTCATTCGCGG	271
DB	62	TCGAGGTTCGAGGTGCGACTCGACGACGGGACGATTGCCGGGGGCTGTCCCGAGTGGG	121
QY	272	CATCCACCGCGCTCCAGAGGCTCATGAGTGGGTGACGGTGGCGATCGCTACTCGGCA	331
DB	122	CGTCCACCGTTCAGTTCGAGCGGTTCGAACTGCGGATGTTGGGACCGTTACGGCGTA	181
QY	332	AGGGCGTTTGAAGGCGATTGAAACGTCACGAAAGAAATCGGCGACGAGTTCGCTGGCC	391
DB	182	AGGGCGTGGAGAAAGGCGTTCGCGCGGTCAACGAGGAAATCTCCGAAACAGATCGTGGT	241
QY	392	TAGAGCTGACGATCAGCGCTCATCGACGAGCTCATGAGTGGGTGACGATCAAGCTTGATGGC	451
DB	242	ACGAGCGGAGGACGCGCTCGTGCAGAGCGCTCGTGCAGAGCGCTCGATCACCCCTGGACG	301
QY	452	ACAAGTCCCGCTGGGTGCAACCGCAATCCTTGGTGTTCATGGCTGTTCGAAAGGCTG	511
DB	302	ACAAGTCCCGCTGGGCGCAACGCCATCTTGGGTGTCTCCCTGGCGGTGGCCAGGCGG	361
QY	512	CTGCTGATTCGCGAGGCTCCCACTGTTCGCTGATCATCGGTGACGAAACGCAACAGTTC	571
DB	362	CCGCGGAGAGCGGACCTTCGCGTCTTCGCTACCTTGGCGCGCCGAAACGCGCACGTGC	421
QY	572	TTCCAGTTCGAATGATGAACATCATCAACGGTGGCGCTCACGCTGACTCGGTGTTGACG	631
DB	422	TGCGGTGCGATGATGAACATCTCAACGGGGGTGGCAGCGGACCAACAGTGCACA	481
QY	632	TTCAGGAATTCATGATCGCTTCAATCGGTGACAGACCTTCTCTGAGGCTTCCGCAAG	691
DB	482	TCCAGGATTCATGATCGCCCCCATCGGGCGGAGAGCTTCCGGAGGCGCTCGCTGGG	541
QY	692	GCGCGAGGTTCACGAGCACTGAAGTCCGTGATCAAGGAAGAGGCGCTTCCACCGGAC	751
DB	542	GCGCGAGGTTCACCACTCGCTCAAGCGGTGCTCAAGGCCCAACGCGCTGGCCACCGGTG	601
QY	752	TTGGCGATGAGGCGGCTTCGCTTCCTTCGTCGGCTCCACCGTGAAGGCTCTTGACTTA	811
DB	602	TCGGTACAGAGGCGGTTTCGCGCCCAACCTGACAGCAACCGCGCGCTGACCTGA	661
QY	812	TCGTTAAGCAATTCGAAAGGCTGGCTTCAACCCAGGCAAGGACATCGCTCTTGTCTGG	871

RESULT 13

US-10-282-122A-25949	
; Sequence 25949, Application US/10282122A	
; Publication No. US20040029129A1	
; GENERAL INFORMATION:	
; APPLICANT: Wang, Liangsu	
; APPLICANT: Zamudio, Carlos	
; APPLICANT: Malone, Cheryl	
; APPLICANT: Haselbeck, Robert	
; APPLICANT: Ohlsen, Kari	
; APPLICANT: Zyskind, Judith	
; APPLICANT: Wall, Daniel	
; APPLICANT: Trawick, John	
; APPLICANT: Carr, Grant	
; APPLICANT: Yamamoto, Robert	
; APPLICANT: Forsyth, R.	
; APPLICANT: Xu, H.	
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms	
; FILE REFERENCE: ELITRA.034A	
; CURRENT APPLICATION NUMBER: US/10/282,122A	
; CURRENT FILING DATE: 2003-02-20	
; PRIOR APPLICATION NUMBER: 60/191,078	
; PRIOR FILING DATE: 2000-03-21	
; PRIOR APPLICATION NUMBER: 60/206,848	
; PRIOR FILING DATE: 2000-05-23	
; PRIOR APPLICATION NUMBER: 60/207,727	
; PRIOR FILING DATE: 2000-05-26	
; PRIOR APPLICATION NUMBER: 60/230,335	
; PRIOR FILING DATE: 2000-09-06	

DB	662	TCAGCGAAGCCATCACGAAAGCGGGTTCACGCTGGGCCAGGACATCGCCCTGGCCTTG	721
QY	872	ACGTTGCTCTCTCTGAGTTCTTCAAGGAGGCGACCTACCACTTCGAAGGTGGCGAGCACT	931
DB	722	ACGTGGCCGCCACCGAGTTCTACGCGGACGGCGCTTACCAAGTTTCGAGGGCTCCAGCGCA	781
QY	932	CGCAGCTGAGATGGCAAAAGCTTTACGCTGAGTTCGTTGACGCGTACCCCAATCGTCTCCA	991
DB	782	GCSCGAGGAGATGGCGCGCTACTACACGAGCTGGTTCGAGTCTCTACCCGCTGGTTCGA	841
QY	992	TCGAGGACCCACTTCGAGGAGATGACTGGGAGGTTTACACCACTTCACCGCAACCATCG	1051
DB	842	TCGAGGACCCGCTCAGCGAGGAAGACTGGCGGGGTGGAAGGCGCTCACCGAGTCCCTCG	901
QY	1052	GGCACAAGGTTTCAGATCGTTGGCGAGACTTCTTCGTCACCAACCTTGAGCGCTGAAGG	1111
DB	902	GGGACCGGCTGCAACTGGTTCGGCGACGACTGTTCTGTCACCAACCCGAGCGGCTGCAGC	961
QY	1112	AGGGCATCGTTAAGAAAGCTGCCAACTCCATCTCTGTTAAGGTGAACCGATCGGTACCC	1171
DB	962	GCGGTATCGACGAGGCGCTGCCAACTCGTCTGCTCAAGGTGAACCAAGATCGGTACGC	1021
QY	1172	TCACCGAGACCTTTCGACGCTGTCGACATGSGCTCACCGCGCAGCTACCTCCATGATGT	1231
DB	1022	TCAGCGAGACTTTTGGACGCGGTGCTCGTGGCGAGCGCAACCGGTACACCGCGATGATCA	1081
QY	1232	CCCACCGTTCGCTGAGACCGAGGACACCACTTCTGACCTTCGCGAGTTTGCACTCAACT	1291
DB	1082	GCCACCGTCTGCTGAGACCGAGGACACCACTGCGCGACATCGCGGTGGCGACGAACT	1141
QY	1292	GTGGCCAGATCAAGACTGCTGCTCAGCAGCTTCGACCGTTCGCAAGGTGCAAGTACAACAGC	1351
DB	1142	CGGGGCGAGATCAAGACGGGTGCCCGCGCGCAGCGAGCGGGTTCGCAAGTACAACAGC	1201
QY	1352	TTCTCCGATCGAGCAGCTGCTTGGCGACGCGCGGTCTACGAGGTTCGCGAGCTCCAGCGCATTC	1411
DB	1202	TGCTCGGATCGAAGAGAGCTCGACGACGAGCGCTCTACGCGGAGCAACGCGGTTC	1261
QY	1412	CACGTTT 1418	
DB	1262	CGCGTTT 1268	

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25949
; LENGTH: 1287
; TYPE: DNA
; ORGANISM: Mycobacterium avium
US-10-282-122A-25949

Query Match      41.0%; Score 647; DB 17; Length 1287;
Best Local Similarity 69.9%; Pred. No. 2.6e-188;
Matches 888; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

Qy 151 GTGGCTGAATCATGCAAGTATTGCTTCGCGAAATTCGACTCCCGCGGTAAACCAACC 210
Db 1 GTCCGGATTATCGAGCAGGTGCGGGCCCGGAGATCTCGACTCCCGCGGTAAACCGACA 60

Qy 211 GTCAGGACAGAGTTTTCCTGGATGACGGTTCCCAACGGTGTGCGAGTGTTCATCCGGC 270
Db 61 GTCAGGTCGAGATCGCCCTGACCGGAACGTTCCCGCGCGCGCGTCCCGTCCGGG 120

Qy 271 GCATCACCGCGCTCCACGAGGCTCATGAGTGGTGCACGGTGGCGATCGCTACCTGGGC 330
Db 121 GCGTCGACCGGTGAGCAGAGGCTGTCTGAATCGCGCACCGCGGGAGCGGTACGGCGC 180

Qy 331 AAGGGCGTTTTGAAGGCGAGTTGAAACGTCACGAAGAAATCGGCGACGAGCTCGCTGGC 390
Db 181 AAGGGCGTGCAGAAAGGCGGTGCAGCGCTGCTGACAGATCGGCCGCGGTGATCGGG 240

Qy 391 CTAGAGGCTGACGATCAGCGCTCATCGAAGAAATGATCAAGCTTGATGGACCGCC 450
Db 241 CTGAACCGCCGACGACGAGCGCTGTTGGAACAGGCGCTCGTGCACCTGGAGCGGACCGCG 300

Qy 451 ACAAAGTCCGCGCTGGTGCAGAGCAATCTTGGTGTTCATGGCTGTTGCAAGGCT 510
Db 301 GACAAGTCGAGGCTGGGCGGCAACGCGATCTTGGCGGTGCTGCGCGGTGGCCAGGCG 360

Qy 511 GCTGCTGATTCCGCGAGCGCTCCCACTGTTCCGCTACATCGGTGACCAAAACGACACGTT 570
Db 361 GCGCGGATTCCGCGGAGTTGCGCGTGTTCGCTACCTGGCGGGGCCCAAGCGGCATC 420

Qy 571 CTTCAGTTCCTAATGATGAACATCATCAACGGTGGCGCTCAACGCTGATCTCGGTGTTGAC 630
Db 421 CTGCGGTGCGATGATGAACATCTCAACGGCGGTGCGACGCGGACACCGCGCTGAC 480

Qy 631 GTTCAGAAATTCATGATCGCTCCATCGGTGCGAGACCTTCTTGAGGCTCTCCGCAAC 690
Db 481 ATCCAGAGTTCATGTTGCGCGCCGATCGCGCGCGCGAGTTTCGCGGAGGCAATTGCGCTG 540

Qy 691 GCGCGGAGGTCTACCAACGCACTGAAGTCCGTATCAAGGAAAGGGCCGTGCCACCGGA 750
Db 541 GTTGGCGAGGTGATCTACTGCTGAAGTGGTGTGCTGAAGAGGGGCCCTTAGCACCGC 600

Qy 751 CTTGGGATAGGCGCGCTTCTGCTCTCTCCGTCCGCTCCACCGGTGAGGCTCTTGACCTT 810
Db 601 CTGGGCGACGAGGCGCGCTTCTGCGCCCGACGCTGGCGGCGCACCAACCGCGCTGACCTG 660

Qy 811 ATCGTTAAGCAATCGAGAAAGGTGGCTTACCCCGAGGCAAGGACATCGCTTCTGCTCTG 870
Db 661 ATCGGCGGCGCCATCGAATCGGCGCGGCTTCAAACTCGGCGACCGAGCTGGCGCTGGCCCTC 720

871 GAGCTTGCTTCTCTGAGTTCTTCAAGGACGGCAC-----CTACCACCTTCGAGGTGGCCAG 927
872 GACGCGCGCGCCACCGAGTTCTACAGCGACGGCACCGGCTACAGTTTCGAGGGCGAGCACC 780
928 CACTCCGAGCTGAGATGGCAAAAGTTTACGCTGAGCTGTTGACGGGTACCCCAATCGTC 987
781 CGACCGCGCGAGCAGATGGCGCGAGTTCTACGCGCGGCTGCTCGCGCGGTATCCCGTTGGTG 840
988 TCCATCGAGGACCCACCTGCAAGGAGATGACTGGGAGGGTTACACCAACCTTCACCGCAACC 1047
841 TCCATCGAAGACCCGCTGTCGGAAGATGATGGAGCGCTGGGCGGCGCTGACCGCGTCG 900
1048 ATCGGCGCAAGGTTTCAGATCGTTGGCGACGACTTCTTCGTACCAACCTTCGAGCGCTG 1107
901 ATCGGCGACCGGTTGAGCTCGTCGCGGACGAGCTCTTCGTACGAAACCCCGAAGCGCTG 960
1108 AAGAGGGGATCGTAAGAAGGCTGCCAATCTCATCTCTGTTAGGTTGAACCAAGATCGGT 1167
961 GAAGAGGGTATCGAGAAGGGCGTCGCAAAATGCGTTGCTGTCGAAGGTGAATTCAGATCGGC 1020
1168 ACCTTCACCGAGACCTTCGAGCGTGTGACATGGCTCACCGCGCAGGCTACACCTCCATG 1227
1021 ACCTGACCGAGAGCGTGGAGCGCTGCGCTGGCCACACACGCGGTACCGGACCATG 1080
1228 ATGTCGCCACCGTTCCGCTGAGACCGAGGACACCAACATTCGTCACCTTCGAGTTGCACTC 1287
1081 ATGAGCCACCGCAGCGGTGAACCGAGGACACCAACGATCGCGACCTTGGCGGTGCGCGTC 1140
1288 AACTGTGGCGCAGATCAAGACTGGTGTCTCAGAGAGTTTCGACCGTGTTCGCAAGTACAAAC 1347
1141 GGCAGCGGCGAGATCAAGACCGGAGCGCGCGCCGCGAGCGGGTGGCCAAATACAAAC 1200
1348 CAGCTTCTCCGATCGAGCAGCTGCTGGCGAGCGCGCGCTCTACGAGGTTCGAGCGCA 1407
1201 CAGCTGTGCGGATCGAGGAGGCGCTCGCGGACCGCGCGCTTACGCGCGGACCTGGCC 1260
1408 TTCCCGCGCTT 1418
1261 TTCCCGCGGTT 1271

RESULT 14
US-10-080-170-649
; Sequence 649, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: CODE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495, 0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 649
; LENGTH: 45191
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-649

Query Match      40.1%; Score 632.6; DB 15; Length 45191;
Best Local Similarity 68.4%; Pred. No. 3e-183;
Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

Qy 140 AGGAGGCCACAGTGGCTGAAATCATGACGCTATTTCGCTCGCAAAATCTCGACTCCCGCG 199
Db 19931 AGGAGAACCCAGTGGCGGATTCGAGCAGGTAGGCGCCCGAGAGATCCTCGATTCCCGCG 19990
200 GTAACCCCAACCGTCGAGGCGAGAGGTTTCTCTGGATGACGGTTCCCAAGGTTGTCGAGGTG 259
```


19991 GCAACCCGACGGTGGAGGTCGAGGTGGCGCTTATCGACGGGACAAATTGCGCCGGCGCGG 20050
260 TTCCATCCGGCGCATCCACCGCGTCCACGAGGCTCATGAGCTCGGTGCGGCAATC 319
20051 TGGCGTCCGGCGCGCTCCACCGGGAGACGAGGCGTTCGAGTTGCGGACGGCGCGCATC 20110
320 GCTACCTGGGCAAGGGCGTTTGAAGGCAGTTGAAGAACGTCACGAAAGAAATCGGCGACG 379
20111 GCTACGGCGCAAGGGTGCAGAAAGCCTGTCAGGCTGTTCTTGATGAGATCGGCCGG 20170
380 AGCTCGTGGCTAGAGGCTGACGATCAGCCCTCATCGACGAAGCAATGATCAAGCTTG 439
20171 CCGTCATCGGACTCAACCGCGACGACGAGCTGCGTGTGTTCCGTTATGTCGGGGGCGCA 20230
440 ATGGCACCGCAACAGTCCGGCTGGTGGTCAACGCAATCCTTGGTGTTCATGCTG 499
20231 ACGCACCCCGCAAGTCCGGCTGGCGGCGCAACGCGATCTTGGGTGCTCGCTCGCTG 20290
500 TTGCAAGGCTGTGCTGATTCGCGAGGCTCCACACTGTTCCGCTACATCGGTGGACCA 559
20291 TTGCCAAGGCGCGCGGATTCGGCGAGCTGCGTGTGTTCCGTTATGTCGGGGGCGCA 20350
560 ACGCACAGTCTTCCAGTTCCAAATGATGAACATCAACGCTGGCGCTCAGCTGACT 619
20351 ACGCGACATCTTCCGCTACCGGATGATGAATCTCTCAACGGCGCGCACACGCCGATA 20410
620 CCGGTGTGAGCTTCAGGAATTCATGATCGCTCAATCGGTGAGAGACCTTCTTGAGG 679
20411 CCGCTGTCGACATTCAGAGATTCATGTTGCGGCGCAATTTGGCGCGCCAGCTTCGTGAGG 20470
680 CTCTCCGCAACGGCGCGGAGGCTTACCACGACTGAAAGTCCGTCATCAAGGAAAGGGCC 739
20471 CGTTGCGCTGGGCGCTGAGTGTACCGCGCTCAAGTCCGTCCTGCTGAAGAGAGGGC 20530
740 TGTCCACCGGACTTGGCGATGAGGGCGGCTTCCGCTTCCGTCGCTCCACCGTGAAG 799
20531 TGTCCACCGGCTGGGCGACGAAGGGCGCTTCGCGCCCGATGTGGCGCGCACACCGCGG 20590
800 CTCTTGACCTTATCGTTAAGGCAATCGAAGGCTGCTTCAACCGGCGGCTTTCGAGTCCG 859
20591 CGTTGACCTTGATCAGCGGGCCATCGAGTCGGCGGCTTTCGACCGGGCGCGACGTTGG 20650
860 CTCTTGCTTGGAGTTGCTTCTCTGAGTTCTTCAAGGACGGCAC---CTACCACTTCG 916
20651 CGCTGGCCCTGGACGGCGGCGCACCGAGTTCTTCCGACGCGACCGGCTAGTCTTCG 20710
917 AAGTGGCCAGCACTCGGAGCTGAGATGCAAGGTTACGTGCTGAGCTGTTGACGCGT 976
20711 AGGGCACCAACCGTACCGCAGACCATGATGACCGAGTTCTACGGGGCGCTGCTCGGCGCT 20770
977 ACCCAATCGTCTCATCGAGGCCCACTGCGAAGATGACTGGGAGGTTACCAAC 1036
20771 ACCCGTGTGTGTCATCGAAGACCCCACTGTTCGAAGACGATGGGACGCTGGGCGCGC 20830
1037 TCACCGCAACCATCGGCGCAAGAGTTTCAGATCGTTGGCGACGACTTCTTCGTCACCAAC 1096
20831 TGACGGCTCGATCGGTGACCGGTCGAATCTCGGCGACGACATCTTGTGACCAATC 20890
1097 CTGAGGCTGAAGAGGGGATCGCTAAGAGGCTGCAACTCCATCTCTGTTAAGTGA 1156
20891 CCGAGCGGCTCGAGGAGGGCATCGAACCGGGCGCTGGCAATGCTTGTGCTCAAGGTGA 20950
1157 ACCAGATCGGTACCTTCACCGACCTTCGACGCTGTGACATGGCTCAACCGGAGGCT 1216
20951 ACCAGATCGGAGCTTGAACGAGACATTCGACCGGTTCAGCTGGCTTCACCGCGGAT 21010
1217 ACACCTTCATGATGCCCGTTCGGTGAGACCGGAGCACCAACATTCGTGACTCG 1276
21011 ACGCACGATGATCAGTCAACCGGAGTGGCGAGACGAGGACACCATGATCGCGGACTCG 21070
1277 CAGTTGCATCACTGTGGCGCATCAAGCTGCTGCTCAGACAGGTTCCGACCGTGTG 1336
21071 CCGTGGCCATCGCGACGGGCGAGATCAAGACGGGCGCGCTGCTCGCAGTGAAGCGGCTG 21130

QY 1337 CAAAGTAAACACGACTTCCGCAATCGAGCAGTGTCTTGGCGACGGCGGCTTACGCAG 1396
Db 21131 CAAAATACAAACGAGCTGCTCGGATCGAAGAGCGCTTGGCGACGGCGCCGCTACGCG 21190
QY 1397 GTCGACGCGATTCACACGCTTTCAGGGCTTAATAAAGCGCT 1439
Db 21191 GCGACTGGCAATTCCTCGGTTCCGCTGCGAGACGAAATAGGT 21233

RESULT 15

US-10-080-170-649
; Sequence 649, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 649
; LENGTH: 45191
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-649

Query Match 40.1%; Score 632.6; DB 19; Length 45191;

Best Local Similarity 68.4%; Pred. No. 3e-183;

Matches 891; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

QY 140 AGGAGGCCACAGTGGCTGAAATCATGACGTATTTCGCTCGGAAATTCGACTCCGCG 199
Db 19331 AGAGAAACCAAGTCCCGATTCGACGAGGTAGGGCCCGAGAGATCTTCGATTCGCG 19990
QY 200 GTAACCCAAACCGTCGAGGCGAGAGTTTTCCTGGATGACGCTTCCCAACGCTGTCGACGGTG 259
Db 19991 GCAACCCGACGGTGGAGTTCGAGTGGCGCTTATCGACGGGACATTCGCCCGGCGCG 20050
QY 260 TTCATCCCGCGCATCCACGGGCTCCAGAGCTCATGAGTGCCTGAGCGTGGCGATC 319
Db 20051 TGGCGTCGGCGCTCGACCGGGGAGCACGAGGCTGCGAGTTCGCGACGGCGGATC 20110
QY 320 GCTACCTGGGCAAGGGCGTTTGAAGGCGAGTTGMAAACCTCAACGAAGAAATCGCGGACG 379
Db 20111 GCTACGGCGCAAGCGGTGCAAAAAGCCGTCGAGGCTGTTCTTGATGAGATCGGCCCGG 20170
QY 380 AGCTCGCTGGCTTAGAGGCTGACGATCAGCGCTCATCGACGAAGCAATGATCAAGCTTG 439
Db 20171 CCGTCATCGGACTCAACCGCGACGACCGAGTTCGTCGACCGGCGCTGGTGGACCTAG 20230
QY 440 ATGGCACCGCAACAGTCCCGCTGGGTGCAAAACGAATCCTTGGTGTTCATGGCTG 499
Db 20231 ACGGACCCCGCACAGTCCCGCTGGGCGCAACGCGATCTTGGGTGCTCGCTCGCTG 20290
QY 500 TTGCAAGGCTGCTGCTGATTCGCGAGGCTCCCACTGTTCCGCTACATCGGTGGACCAA 559
Db 20291 TTGCCAAGGCGGCGGATTCGGCGAGCTCGCTGTTGTTCCGTTATGTCGGGGGCGCAA 20350
QY 560 ACGCACAGCTTCTTCAGTTCCAATGATGAACATCATCAACGCTGGCGCTCAGCTGACT 619
Db 20351 ACGCGCACATCTGCGGTAACCGATGATGAACATCTCAACGGCGCGCACACGCCGATA 20410
QY 620 CCGGTGTGACCTTCAGGAATTCATGATCGCTCCAATCGGTGAGAGACCTTCTTGAGG 679
Db 20411 CCGCTGTGACATTCAGAGATTCAGGTGGCGCCAAATTTGGCGCGCCAGCTTCGTGAGG 20470
QY 680 CTCTCCGCAACGGCGCGGAGGTCCTACCGACTGAAAGTCCGTCATCAAGGAAAGGGCC 739

```
Db      20471  CGTTGGCGCTGGGGCGCTGAGGTGTACCAACGGCTCAAGTCGGTCTCAAAAAAGGAGGGGC 20530
Qy      740    TGTCCACCGGACTTGGCGGTAGAGGGGGCTTCGTCTCTCCGTCCGTCCGCTCCACCGCTCAGG 799
Db      20531  TGTCCACCGGCGCTGGGGCGGAGAGGGCGCTTCGCCCGGATGTGGCCGGCACCACCGCGG 20590
Qy      800    CTCTTGACCTTATCGTTAAAGCAATCGAAGGCTGGCTTCACCCAGGCAAGGACATCG 859
Db      20591  CGTTGGACCTGATCAGCCGGGCCATCGAGTCGGCGGCTTTCGACCCCGCGCGACGTGG 20650
Qy      860    CTCTTGCTCTGGAGGTTCCTCTCTGAGTCTTCAAGGACGGAC---CTACCACTTCG 916
Db      20651  CGCTGGCCCTTGGACGGCGCGGCCACCGAGTTCTTACCGACGGCACCGGCTAGCTCTTCG 20710
Qy      917    AAGGTGGCCAGCACTCCGCACTGAGATGCAACAGTTTACGCTGAGCTCGTTGACGCGT 976
Db      20711  AGGGACCAACCGGTACCGCAGACGAGATGACCGAGTTCTACGGCGGCTCTCTCGGCGCT 20770
Qy      977    ACCCAATCGTCTCCATCGAGGACCCACTGCAGGAAGATGATGGAGGGTTACACCAACC 1036
Db      20771  ACCCGTGTGTGTCGATCGAAGACCCACTGTCCGAAGACGATTGGACGGCTGGGCCCGCG 20830
Qy      1037  TCACCGCAACCATCGGCGCAAGGTTTCAGATCGTTGGCGAGGACTTCTCGTCACCAACC 1096
Db      20831  TGACGGCTCGATCGGTGACCGGGTGCAATCGTGGCGACGACATCTTTGTCAACCAATC 20890
Qy      1097  CTGAGGCGCTGAAGGAGGGCATCGCTAAGAAAGCTGCCAACTCCATCCCTCGTTAAGGTGA 1156
Db      20891  CCGAGCGGCTCGAGGAGGGGCATCGAACGGGGCGTGGCAATCGTTGCTCGTCAAGGTGA 20950
Qy      1157  ACCAGATCGGTAGCCCTCACCGAGACCTTTCAGACGTGTGACATGGCTCACCGCGCAGGCT 1216
Db      20951  ACCAGATCGGACGTTGACCGAGACACTCGACGGGTTCACGCTGGCTCACACGGCGGAT 21010
Qy      1217  ACACCTCCATGATGTCCCAACGTTTCGGTGGAGACCGGAGGACACCAATTTGCTGACCTCG 1276
Db      21011  ACCGCAAGATGATCAGTCAAGTCCGCGAGTGGCGAGACGGAGGACACCATGATCGCCGACCTCG 21070
Qy      1277  CAGTTGCACTCAACTGTGGCCAGATCAAGACTGGTGTCTCCAGCACGTTCCGACCGGTGCG 1336
Db      21071  CGGTGGCCATCGGACGGGCGAGATCAAGACGGGCGGCTCTCGCAGTGAAGCGGTGCG 21130
Qy      1337  CAAAGTACAAACAGCTTCTCCGCATCGAGCAGCTGTTCGGCAGCGCGGCTCTACGCGAG 1396
Db      21131  CAAATACACACAGCTCTCGGATCGAAGAGGCGCTTGGCGAGCGGCGCGCTACGCGG 21190
Qy      1397  GTCGACGCGCATTTCCACGCTTTTCAGGGCTTAATAAAAGCGCT 1439
Db      21191  GCGACCTGGCATTTCTCTCGGTTCGCGTGGCGAGACGAAATAGGT 21233
```

Search completed: September 29, 2005, 12:00:10
Job time : 4259.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 02:56:03 ; Search time 28.5 Seconds
(without alignments)
1434.811 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VAEIMVFAREILDSRGNPT.....QLLGDAGVYAGRSAPRPFQ 425

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553.5	72.1	429	2 B70623	probable enolase -
2	1502.5	69.7	447	2 G86940	probable enolase [
3	1492.5	66.0	424	2 A97333	enolase (2-phospho
4	1412.5	65.5	426	2 H87462	enolase [imported]
5	1402.5	65.1	429	2 D84094	enolase (2-phospho
6	1391.5	64.5	425	2 A43358	phosphopyruvate hy
7	1377.5	63.9	430	2 B69620	phosphopyruvate hy
8	1359.5	63.1	430	2 AG1381	enolase homolog en
9	1359.5	63.1	430	2 AH1750	enolase homolog en
10	1354.5	62.9	431	2 G96987	enolase [imported]
11	1334.5	61.9	426	2 F70343	enolase - Aquifex
12	1322.5	61.3	422	2 D75251	enolase - Deinococ
13	1320.5	61.3	432	2 S76837	phosphopyruvate hy
14	1318.5	61.2	393	2 AB2752	enolase [imported]
15	1313.5	61.0	430	2 H75022	phosphopyruvate hy
16	1301.5	60.4	428	2 AB1841	phosphopyruvate hy
17	1301.5	60.4	429	2 AC2248	enolase [imported]
18	1301.5	60.4	428	2 A45732	phosphopyruvate hy
19	1298.5	60.3	428	2 D81100	enolase NMB1285 [i
20	1297.5	60.2	428	2 F71209	probable phosphogl
21	1283.5	59.6	429	2 G72323	phosphopyruvate hy
22	1280.5	59.4	434	2 T47276	phosphopyruvate hy
23	1279.5	59.4	434	2 A89851	enolase [imported]
24	1278.5	59.3	429	2 H83191	enolase PA3635 [im
25	1272.5	59.0	433	2 D86705	phosphopyruvate hy
26	1258.5	58.4	434	2 D98001	phosphopyruvate hy
27	1256.5	58.3	434	2 E95130	enolase [imported]
28	1249.5	58.0	433	2 H70141	enolase (eno) homo
29	1241.5	57.6	427	2 A64329	phosphopyruvate hy

30	1232.5	57.2	432	2 F71278	probable enolase (
31	1229.5	57.0	432	1 NOEC	phosphopyruvate hy
32	1229.5	57.0	432	2 G91083	enolase [imported]
33	1226.5	56.9	432	2 H85928	enolase [imported]
34	1224.5	56.8	432	2 AC0859	enolase [imported]
35	1222.5	56.7	431	2 AB0410	phosphopyruvate hy
36	1218.5	56.5	436	2 E64103	phosphopyruvate hy
37	1211.5	56.2	461	2 F82700	enolase XF1291 [im
38	1204.5	55.9	422	2 D86659	phosphopyruvate hy
39	1178.5	54.7	414	2 E81264	phosphopyruvate hy
40	1174.5	54.5	433	2 E82076	enolase VC2447 [im
41	1168.5	54.2	428	2 A72034	enolase - Chlamydo
42	1165.5	54.1	428	2 G81504	enolase CP1071 [im
43	1165.5	54.1	428	2 F86590	enolase [imported]
44	1155.5	53.6	424	2 E81654	enolase TC0876 [im
45	1155.5	53.6	444	2 T12341	phosphopyruvate hy

ALIGNMENTS

RESULT 1

B70623

Probable enolase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 16-Aug-2004

C:Accession: B70623

R:Coile, S.T.; Brosch, R.; Garnier, J.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

R:Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70623

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-429 <COL>

A:Cross-references: UNIPROT:P96377; GB:292539; GB:AL123456; NID:g3261714; PIDN:CAB06856.

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: eno

C:Superfamily: Enolase

Query Match	72.1%	Score 1553.5;	DB 2;	Length 429;
Best Local Similarity	73.2%	Pred. No. 1.3e-93;		
Matches 308;	Conservative 42;	Mismatches 70;	Indels 1;	Gaps 1;

QY 4 IMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLKGV 63

DB 4 IEQVRAEILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYLKGV 63

QY 64 LKAVENNEIEIGDELAGEADQRLIDEAMIKDGTANKSRIGANAILGVSMVAKAAAD 123

DB 64 QKAVQALVEIGFAVIGLNADDQRLVDQALVDLDGTPDKSRGLGNAILGVSLVAKAAAD 123

QY 124 SAGLPLFRYVIGPNNAHLVPPMNIINGGAHDSGVDOEFMIAPICAEFTSALRNGAE 183

DB 124 SAEPLFRYVIGPNNAHLVPPMNIINGGAHDSGVDOEFMIAPICAEFTSALRNGAE 183

QY 184 VYHAKSVIKKEGLSTGLDEGGFAPSVGSTRALDLIVKAIKAGFTPGKDIALLADVA 243

DB 184 VYHAKSVLKKEGLSTGLDEGGFAPDVAGTTAALDLISRAIESAGLRPGADVALALDAA 243

QY 244 SSEFFKQGT-YHIEGQGHSAEAMNVAELVDAYPIVSIEDPLQEDDWBGYTWLTATIGD 302

DB 244 ATEFFTDGTYVFEQTRTADQMTFYGALLGAYPLVSIEDPLSDDDWGDGMAALTASIGD 303

QY 303 KVOIVGDDDFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHACGYTSMWSH 362

DB 304 RVOIVGDDDFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHACGYTSMWSH 363

QY 363 RSGETEDTTIADLAVALNCCQIKTGAPARSDRVAKYNQLLRIBQLLGDAGVYAGRSAPFR 422

Db 364 RSGTETMTADLAVALGSGQIKTGAPARSERVAKYNQLRIEALGDAAARYAGDLAFPR 423
Qy 423 F 423
Db 424 F 424

RESULT 2
G86940
probable enolase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: G86940
R.; Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
e, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: G86940
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>
A:Cross-references: GB:AL450380; NID:gl3092594; PIDN:CAC29763.1; GSPDB:GN00147
C:Genetics:
A:Gene: eno
C:Superfamily: enolase

Query Match 69.7%; Score 1502.5; DB 2; Length 447;
Best Local Similarity 70.8%; Pred. No. 2.8e-90;
Matches 300; Conservative 47; Mismatches 76; Indels 1; Gaps 1;

Qy 1 VAEMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVVHAEHLRDGGRYL 60
Db 19 VPVIEQVAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVVHAEHLRDGGRY 78
Qy 61 KGVKAVENVNEEIGDELAGLEADDORLIDEAMIKLDGTANKSRILGNATILGYSMAVAKA 120
Db 79 KGVRAVDVLDIEGPVIGINANDOKRLIDQELLDLDTGTPDKSRILGNATILGYSMAVAKA 138
Qy 121 AADSAGLPLFRYIGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPICGAETFSBALRN 180
Db 139 AADSAGLPLFRYIGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPICGAETFSBALRN 198
Qy 181 GARVYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAEKAGTTPGKDIALLAL 240
Db 199 GARVYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAEKAGTTPGKDIALLAL 258
Qy 241 DVASSEFFKDGTYHPEGGGSHAAEMANVYAEVLDVAPIVSIEDPLQEDDMEGYTNLTAT 299
Db 259 DAAATFYTDCIGHYHEGTHADQWTEFTADLGSTPLVSIEDPLQEDDMEGYTNLTAT 318
Qy 300 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHAGYTSM 359
Db 319 IGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHAGYTSM 378
Qy 360 MSHRSGETEDTTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEALGDAGVYAGRA 419
Db 379 ISHRSGETEDTTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEALGDAGVYAGRA 438
Qy 420 FPRF 423
Db 439 FLRY 442

RESULT 3
A97533
enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) [imported]
C:Species: Acrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 16-Aug-2004
C:Accession: A97533

R.; Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A97533
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <KUR>
A:Cross-references: UNIPROT:Q8UFH1; GB:AE007869; PIDN:AAK87218.1; PID:gl5156500; GSPDB:G
C:Genetics:
A:Gene: AGR_C_2631
A:Map position: circular chromosome
C:Superfamily: Enolase

Query Match 66.0%; Score 1422; DB 2; Length 424;
Best Local Similarity 66.2%; Pred. No. 4.5e-85;
Matches 276; Conservative 58; Mismatches 81; Indels 2; Gaps 1;

Qy 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGGRYLKGV 63
Db 4 ITDIIAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGGRYLKGV 63
Qy 64 LKAVENVNEEIGDELAGLEADDORLIDEAMIKLDGTANKSRILGNATILGYSMAVAKAAD 123
Db 64 EKAVEAVTEIFDAICGFDAENQIQIDQMIALDGTNPKSRILGNATILGYSMAVAKAAD 123
Qy 124 SAGLPLFRYIGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPICGAETFSBALNGAE 183
Db 124 ASGLPLRYVGGNNAHLVPVPMNIIINGGAHDSGVVDVQEFMIAPICGAETFSBALNGAE 183
Qy 184 VYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAEKAGTTPGKDIALLDVA 243
Db 184 VYHALKSVIKEKGLSTGLDEGGFAPSGVSTREALDLIVKAEKAGTTPGKDIALLDVA 243
Qy 244 SFEFFKDGTYHPEGGGSHAAEMANVYAEVLDVAPIVSIEDPLQEDDMEGYTNLTATIG 301
Db 244 STEFFKDGTYHPEGGGSHAAEMANVYAEVLDVAPIVSIEDPLQEDDMEGYTNLTATIG 303
Qy 302 DKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHAGYTSMS 361
Db 304 NKQQLVGGDLFTVNSARLDGKMGVANSILVKVNIQIGTLTETFDVADMAHAGYTSMS 363
Qy 362 HRSGETEDTTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEALGDAGVYAGRS 418
Db 364 HRSGETEDTTADLAVALNCGQIKTGAPARSDRVAKYNQLRIEALGDAGVYAGRS 420

RESULT 4
H87462
enolase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87462
R.; Nierman, W.C.; Deboy, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87462
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-426 <STO>
A:Cross-references: UNIPROT:Q9A7J9; GB:AE005673; NID:gl3423140; PIDN:AAK23700.1; GSPDB:G
C:Genetics:
A:Gene: CC1724
C:Superfamily: enolase

Query Match 65.5%; Score 1412.5; DB 2; Length 426;
Best Local Similarity 66.4%; Pred. No. 1.9e-84;
Matches 280; Conservative 59; Mismatches 80; Indels 3; Gaps 2;

Qy	124	SAGUPLRILGGPNHVLVPPVPMMLINGCAHADSVDVQEFEMIAIIGMABIFSEALKNRGA	18
Qy	124	SAGUPLRILGGPNHVLVPPVPMMLINGCAHADSVDVQEFEMIAIIGMABIFSEALKNRGA	18
Db	124	ASGLPLRYRYVGGTNAHVLVPPVPMNIIINGCAHADNPIDFQEFMILPVGATSTIREAVRYGSE	183
Qy	184	VYHALKVIKEGLSTGLDGGFPAPSVGSTREALDLIVKAIEKAGFPKGDIALALDVA	243
Db	184	VFHTLKKRLKDAGHNTNVGDEGGFAPNLKNAQAALDFIMESTEKAGFPXGEDIALLDCA	243
Qy	244	SSEFFKDGTYHFEQG--HSAAEHNVVAELVDVAPIVSIIEDPLOEDDWEGYTNLTATIG	301
Db	244	ATEFFKDGNYVEGERTRDPKAQKYLAKLASDYPVITIEDGMAEDDWEHGKYITDLIG	303
Qy	302	DKVQIVGDDPFTVNPRLKEGIAKKAANSILVKVNIQIGTLTETFDADVMAHRAGYTSMMS	361
Db	304	NKCOLVGDDLPVTNSARLRDGIIRLGVANSILVKVNIQIGLSETLDAVETAHKAGYTVMS	363
Qy	362	HRSETEDTTIADLAVALNCCQIKITCAPARSDRVAKYNQLRIEQLLDGAGVYAGRSA	419
Db	364	HRSETEDSTIADLAVANCCQIKTGSLARSDRTAKYNQLRIEELDKQARYAGRSA	421

Qy	243	ASSEFF--KDGTYHF--EGGQHSAAEMANVVAELVDAYPIVTSIEDPLQEDDWEGYTNLTA	298
		:	
Db	244	ASSEFYNKEDGKYHLSGGVVKTSAEMVDMVYEELVSKYPIISIEDGLDENDMEGHKLLTE	303
		:	
Qy	299	TTGDGVQIVGGDDFFVTPNERLKEGIAKKAANSILVKVNLQGLTETFTDAVDMHARAGYTS	358
		:	
Db	304	RLGKKVQLVGDDLFTVNTKKLSEGIKNGVGNLSILVKVNLQGLTETFTDAEMAKRAGYTA	363
		:	
Qy	359	MMSHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNQLLRIEQLLGDGAVVYAGRS	418
		:	
Db	364	VISHRSGETEDSTTIADIAVATNAQGIKTGAPSRPTRVAKYNQLLRIEDQLAETAQYHGIN	423
		:	
Qy	419	AF 420	
		:	
Db	424	SP 425	
		:	

RESULT 8
 AGI381
 enolase homolog eno [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
 C/Accession: AGI381
 R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.;
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournon,
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A/Accession: AGI381
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-430 <GLA>
 A;Cross-references: UNIPROT:Q928I3; GB:NC_003210; PIDN:CAD00533.1; PID:g1641199
 A;Experimental source: strain EGD-e
 C:Genetics:
 A;Gene: eno
 C;Superfamily: Enolase

Query Match	63.1%	Score 1359.5	DB 2	Length 430
Best Local Similarity	63.7%	Pred. No. 5.3e-81		
Matches 269	Conservative	60	Mismatches 88	Indels 5
Gaps	3			
QY	4	IMHYFAEILDSRGNPVTEAEVFLDQSHCVAGVPSGASTGVHEAHELDDGD-RYLKGK	62	
Db	4	ITEVFAEVLDSRGNPVTEVEVTEAGAFGRALVPSGASTGVEAVELRDGDKARYLKGK	63	
QY	63	VLKAVENNERIGDELAGLEADQRLIDEMIKLDGTANKSRIGLANAILGVSMVAKAAA	122	
Db	64	VLKAVENVNDIIADKIIGFDVTDQIGDKAMIELDGTNNKGLGANAILGVSLAARAAA	123	
QY	123	DSAGLPIFRYIGGBNAHVLVPPMMNITINGAHADSGVDVQEFMIAPIGAETSEALRNGA	182	
Db	124	DELGVHLYEYLGGVNGKLVPPMMNITINGEHADNNVDVEEFVMPVPVGNPKEALRWGA	183	
QY	183	EWYHALKSVIKEGLSTGLGDGEGFAPSVGSTREALDLVLKAIKAGTPTGKDIALADLV	242	
Db	184	EILHALKAVLKGKLTNVGDGEGFAPNLKSNEEALETTIMQAIKQAGYKPGVEEVKLAMD	243	
QY	243	ASSFFP--KQGTYHF--EGQHSAAEMANYFAELVDAYFIVSIEDPQEDDMEGYTNLTA	298	
Db	244	ASSFFYNRETGKVELGEGVTRTSEBWTVMYEMITKVPILSIEDGLDENDWDGPKLUTE	303	
QY	299	TIGDKVQIVGGDDPPVTPNPERLKEGIAKKAANSTLVKVNQIGLTTFTFDVADMAHRAGYTS	358	
Db	304	RIGRVQIVGGDDLPVNTTKKEGIEKGTANSILIKVNQIGLTTFTDIAIEMAKRAGVTA	363	
QY	359	MMSHRSGETEDTTIADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRS	418	
Db	364	VISHRSGETEDSTIADIVATNAGQIKTGAPTRTDVAKYNOLLRIENLADLAFYHND	423	

Qy 419 AF 420
Db 424 TF 425

RESULT 9
AH1750
enolase homolog eno [imported] - Listeria innocua (strain Clp11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C;Accession: AH1750
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duseurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.N.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A.;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1750
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <GLA>
A;Cross-references: UNIPROT:Q92813; GB:AL592022; PIDN:CAC97776.1; PID:g16415071; GSPDB:G
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: eno
C;Superfamily: Enolase

Query Match 63.1%; Score 1359.5; DB 2; Length 430;
Best Local Similarity 63.7%; Pred. No. 5.3e-81;
Matches 269; Conservative 60; Mismatches 88; Indels 5; Gaps 3;

Qy 4 IMHVFAEILDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVHEAHELRDGGD-RYLKKG 62
Db 4 ITEVAREVLDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVHEAHELRDGGD-RYLKKG 63

Qy 63 VLKAVENNEEIGDELADDDORLIDEAMIKLDTANKSRIGANAILGVSMVAKAAA 122
Db 64 VLKAVENNDIADKIIGFDVTQIGIDKAMIELDTGPNKGLGANAILGVSLAARAAA 123

Qy 123 DSAGLPLFRYIGGPNHVLVPMNMINGGAHDSGVVDQEFMIAPIGAEFTFSEALRNGA 182
Db 124 DELGVHLYEYLGVGKVLVPMNMINGGAHDSGVVDQEFMIAPIGAEFTFSEALRNGA 183

Qy 183 EVYHAKSVLKEGLSTGLDGGFAPSGVSTREALDLIVKAEKAGFTPGKDIALDV 242
Db 184 EILHALKAVLKGKLTGVDGEGFAPNLKNEAELETIMQAIKDAKPGKEVYKLAMD 243

Qy 243 ASSEFF--KDGTYHF--EGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEYTNLTA 298
Db 244 ASSEFFNRETGKYLKGEVTRISEEMVTVYEEMITKYPISIEDGLDENDWGFKLITE 303

Qy 299 TIGDKVQIVGDDPFVTNPERLKEGIAKAANSILVKVNOIGTTLTETFPDVMMAHRAGYTS 358
Db 304 RIGDRVQLVGDLLFVNTTKLEKIEKIANSLIKVNOIGTTLTETLDAEMAKRAGYTA 363

Qy 359 MMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIBOLLGDAGVYAGRS 418
Db 364 VISHRSGETEDTTIADLAVATNAGQIKTGAPTRDRVAKYNQLLRIBOLLADLAHYGND 423

Qy 419 AF 420
Db 424 TF 425

RESULT 10
G96987
enolase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G96987
R;Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G96987
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-431 <KUR>
A;Cross-references: UNIPROT:Q97L52; GB:AE001437; PIDN:AAK78690.1; PID:g15023593; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0713
C;Superfamily: enolase

Query Match 62.9%; Score 1354.5; DB 2; Length 431;
Best Local Similarity 62.5%; Pred. No. 1.1e-80;
Matches 263; Conservative 73; Mismatches 82; Indels 3; Gaps 2;

Qy 3 EIMHVFAEILDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVHEAHELRDGGD-RYLKKG 61
Db 6 EIVDMARQILDSRGNPTVEAEVFLDDGSHGVAGVPFGASTGVHEAHELRDNDKQAYLKG 65

Qy 62 GVLKAVENNEEIGDELADDDORLIDEAMIKLDTANKSRIGANAILGVSMVAKAAA 121
Db 66 SVLNVDVNNVNETTATELIGNVDFDQTLIDQTMLEIDGTENKSKLGNAMLVSLAVAAA 125

Qy 122 ADSAGLPLFRYIGGPNHVLVPMNMINGGAHDSGVVDQEFMIAPIGAEFTFSEALRNG 181
Db 126 AEVLGSLYQYLGGVNAKVLVPMNMINGGAHDSGVVDQEFMIAPIGAEFTFSEALRSC 185

Qy 182 AEVYHAKSVLKEGLSTGLDGGFAPSGVSTREALDLIVKAEKAGFTPGKDIALD 241
Db 186 AEVYHAKSVLKEGLSTGLDGGFAPSGVSTREALDLIVKAEKAGFTPGKDIAFD 245

Qy 242 VASSEFFKDGTYHF--EGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEYTNLTAT 299
Db 246 PASTEFYENGKYLKGEVYVSEEMVTVYEEMITKYPISIEDGMAEEDWGMKLLTDR 305

Qy 300 IGDVQVIGDDPFVTNPERLKEGIAKAANSILVKVNOIGTTLTETFPDVMMAHRAGYTSM 359
Db 306 IGDVQVIGDDPFVTNPERLKEGIAKAANSILVKVNOIGTTLTETFPDVMMAHRAGYTAV 365

Qy 360 MMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIBOLLGDAGVYAGRS 419
Db 366 VSHRSGETEDTTIADLAVAVNAGQIKTGAPARSDRVAKYNQLLRIBELLGEVAFRGLNA 425

Qy 420 F 420
Db 426 F 426

RESULT 11
F70343
enolase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 16-Aug-2004
C;Accession: F70343
R;Beckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70343
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-426 <AQF>
A;Cross-references: UNIPROT:O66778; GB:AE000692; NID:g2983130; PIDN:AAC06738.1; PID:g2983
A;Experimental source: strain VF5
C;Genetics:
A;Gene: eno
C;Superfamily: Enolase

Query Match 61.9%; Score 1334.5; DB 2; Length 426;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 23:34:02 ; Search time 88.5 Seconds
(without alignments)
2459.137 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VAIMHVFAREILDSRGNT.....QLLGDAGVYAGRSAPFRFQG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2148	99.7	425	1 ENO_CORGL	Q8nrs1 corynebacte
2	2025	94.0	425	1 ENO_COREF	Q8fgs7 corynebacte
3	1913	88.8	425	2 Q6NT61	Q8ni61 corynebacte
4	1395.5	74.0	426	1 ENO1_STRCO	Q9fzq3 streptomyc
5	1567.5	72.7	426	2 Q6ADR6	Q6adr6 leifsonia x
6	1560.5	72.4	429	2 Q741U7	Q741u7 mycobacteri
7	1553.5	72.1	429	1 ENO_MYCTU	P96377 mycobacteri
8	1553.5	72.1	429	2 Q7U0U6	Q7u0u6 mycobacteri
9	1550.5	71.9	428	2 Q82HH5	Q82hns streptomyc
10	1505.5	69.9	426	2 Q6AAB8	Q6aab8 propionibac
11	1501.5	69.7	429	1 ENO_MYCLE	Q9cd42 mycobacteri
12	1436	66.6	428	2 Q83H73	Q83h73 tropheryma
13	1436	66.6	446	2 Q83PF7	Q83pf7 tropheryma
14	1422	66.0	424	1 ENO_AGRP5	Q8ufh1 agrobacteri
15	1412.5	65.5	426	1 ENO_CAUCR	Q9a7j9 caulobacter
16	1408	65.3	429	1 ENO_THETN	Q8r967 thermoaer
17	1402	65.1	429	1 ENO_BACHD	Q9k717 bacillus ha
18	1400.5	65.0	429	2 Q84FY9	Q84fy9 methylobact
19	1398	64.9	425	1 ENO_BRUSU	Q8q0c3 brucella su
20	1397	64.8	424	1 ENO_RHIME	Q92q98 rhizobium m
21	1396	64.8	424	1 ENO_RHIL0	Q98mz3 rhizobium l
22	1391	64.5	425	1 ENO_BRUME	Q8yhf0 brucella me
23	1390.5	64.5	430	2 Q65EN2	Q65en2 bacillus li
24	1377.5	63.9	430	1 ENO_BACSU	P37869 bacillus su
25	1376.5	63.9	431	1 ENO_CLOPE	Q8xku4 clostridium
26	1375	63.8	431	2 Q815K8	Q815k8 bacillus ce
27	1373	63.7	431	2 Q81X78	Q81x78 bacillus th
28	1373	63.7	431	2 Q6HBF3	Q6hbf3 bacillus th
29	1372	63.7	432	2 Q7VDY0	Q7vdy0 prochloroco
30	1371	63.6	431	2 Q631M2	Q631m2 bacillus ce
31	1370.5	63.6	427	1 ENO_RALSO	Q8y0b5 ralstonia s

32 1370 63.6 428 2 Q74AR6
33 1362.5 63.2 430 2 Q7U3T1
34 1362 63.2 431 2 Q72XV5
35 1361.5 63.2 430 2 Q67SV9
36 1359.5 63.1 430 1 ENO_LISIN
37 1359.5 63.1 430 1 ENO_LISMO
38 1359.5 63.1 430 2 Q71WX1
39 1356.5 62.9 431 2 Q898R0
40 1355.5 62.9 429 1 ENO_OCEH
41 1355.5 62.9 434 2 Q8GD25
42 1354.5 62.9 431 1 ENO_CLOAB
43 1350 62.6 431 2 Q7V483
44 1349 62.6 430 2 Q7V377
45 1344 62.4 422 2 Q72H85

ALIGNMENTS

RESULT 1

ID_ ENO_CORGL STANDARD; PRT; 425 AA.
AC Q8NRS1;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocuNames=Cgl0974, cgl111;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Busch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins".
RL J. Biotechnol. 104:5-25(2003).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2O).
CC -I- COPACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

EMBL; AP005277; BAB98367.1; --
EMBL; BX927150; CAP19681.1; --
HSSP; P08324; 1E91.


```

DR HAMAP; MF 00318; -, 1.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR Pfam; PF03952; Enolase_N; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; ENOLASE.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT_SITE 154 154 By similarity.
FT METAL 241 241 Magnesium (By similarity).
FT METAL 282 282 Magnesium (By similarity).
FT METAL 309 309 Magnesium (By similarity).
SQ SEQUENCE 425 AA; 44949 MW; 87AABD82C165FD1D CRC64;

Query Match 99.7%; Score 2148; DB 1; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.5e-119;
Matches 423; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYL 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYL 60

QY 61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSM 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSM 120

QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNINGGAHDSGVVDVQEFMIAPIGAETFS 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AADSAGLPLFRYIGGPNNAHVLVPMNMNINGGAHDSGVVDVQEFMIAPIGAETFS 180

QY 181 GAEVYHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIAL 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GAEVYHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIAL 240

QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGY 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGY 300

QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVDMHRA 360
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVDMHRA 360

QY 361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDAGV 420
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDAGV 420

QY 421 PRFQG 425
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 PRFQG 425

RESULT 2
ENO COREF STANDARD; PRT; 425 AA.
AC Q8Q87;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Names; OrderedLocusNames=CE1042;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y8-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=2273752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
```

```

RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RL Genome Res. 13:1572-1579 (2003).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the enolase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR HSPB; AP005217; BAC17852.1; -.
DR HSPB; P08324; 1E91.
DR HAMAP; MF 00318; -, 1.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR Pfam; PF03952; Enolase_N; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; ENOLASE.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT_SITE 154 154 By similarity.
FT METAL 241 241 Magnesium (By similarity).
FT METAL 282 282 Magnesium (By similarity).
FT METAL 309 309 Magnesium (By similarity).
SQ SEQUENCE 425 AA; 54122 MW; 90744F1C427768C9 CRC64;

Query Match 94.0%; Score 2025; DB 1; Length 425;
Best Local Similarity 93.9%; Pred. No. 4.9e-112;
Matches 399; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYL 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MAEIMHVFAREILDNRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGDRYL 60

QY 61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSM 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 KGVLKAVENNVNEEIGDELADDDORLIDEAMIKLDGTANKSRGLGANAILGVSM 120

QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNINGGAHDSGVVDVQEFMIAPIGAETFS 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AADSAGLPLFRYIGGPNNAHVLVPMNMNINGGAHDSGVVDVQEFMIAPIGAETFS 180

QY 181 GAEVYHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIAL 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GAEVYHALKSVIKEKGLSTGLDGGFAPSVGSTREALDLIVKAIKAGTTPGKDIAL 240

QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGY 300
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQDDWEGY 300

QY 301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVDMHRA 360
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GSKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTTFDAVDMHRA 360

QY 361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDAGV 420
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 SHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNOLLRIEQLLDAGV 420

QY 421 PRFQG 425
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 PRFQG 425
```

```

RESULT 3
ID Q6NI61 PRELIMINARY; PRT; 425 AA.
AC Q6NI61;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name: eno; OrderedLocusNames=DIP0917;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22955443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdeno-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whithead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RL diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523(2003).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
DR EMBL; BX248356; CAB9433.1; -.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase C; 1.
DR Pfam; PF03952; Enolase N; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 425 AA; 45029 MW; 85BF477BE7ECDCA6 CRC64;

Query Match 88.8%; Score 1913; DB 2; Length 425;
Best Local Similarity 87.8%; Pred. No. 2.1e-105;
Matches 373; Conservative 23; Mismatches 29; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDSRNPVTEAEVFLDDGSHGVAGVPGSGATGVVHAHELRDGGERYLG 60
DB 1 MADIMHVFAREILDSRNPVTEAEVFLDDGSHGVAGVPGSGATGVVHAHELRDGGERYLG 60
QY 61 KGVLKAVENYNEBIGELAGLEADDQRLIDEAMI KLDGTANKSRGLGANAILGVSMVAKA 120
DB 61 KGVLNANNVNEEIAAIAEADDQRLIDQAMIALDGTENKSRGLGANAILGVSIYAKA 120
QY 121 AADSAGLPLPRYTGPNNAHVLVPVMNMIINGGAHADSGVDVQEFMIAPICAEFTFSEALRN 180
DB 121 AADSAGLPLRYTGGPNNAHVLVPVMNMIINGGAHADSGVDVQEFMIAPICAEFTFSEALRM 180
QY 181 GAEEVYHAKSVIKEKLGSLGDEGGPAPSVGSTREALDLIVKAIKAGFTPGKDTALAL 240
DB 181 GAEEVYHSLKSVIKSLGSLGDEGGPAPSVGSTREALDLIVKAIKAGFKPGADIALAL 240
QY 241 DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLDVAYPIVSIEDPLQEDDWEGYTALTAI 300

```

```

Db 241 DVASSEFFKDGTYHFEGGQHSAAEMAKVYEQIAEYPIVSIEDPLQEDDWEGYTALTAI 300
QY 301 GDKVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGITLTETFDVADMAHRAGYTSMW 360
DB 301 GDKVOIVGDDFFVTNPARLKEGIEKKAANALLVKVNOIGITLTETFDVADLAHRNGYRTMM 360
QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSAF 420
DB 361 SHRSGETEDTTIADLAVALGCGQIKTGAPARSERVAKYNOLLRIEQLLDGAGVYAGRSAF 420
QY 421 PRFQG 425
DB 421 PRFQG 425

RESULT 4
ENOI STRCO STANDARD; PRT; 426 AA.
ID ENOI STRCO
AC Q9F2Q3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Enolase 1 (EC 4.2.1.11) (2-phosphoglycerate dehydratase 1) (2-phospho-
DE D-glycerate hydro-lyase 1).
GN Name: eno1; OrderedLocusNames=SCO3096; ORFNames=SCE41.05c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Crang C.-H., Kieser T., Lark L., Hidalgo J., Hornsby T., Howarth S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Oliver K., O'Neill S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AL939115; CAC09537.1; -.
CC HSSP; Q9NDH8; 10EP.
CC HAMAP; MF_00318; -.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; Enolase C; 1.
CC Pfam; PF03952; Enolase N; 1.
CC PRINTS; PR00148; ENOLASE.
CC ProDom; PD000902; Enolase; 1.
CC TIGRFAMs; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.

```

KW Complete proteome; Glycolysis; Lyase; Magnesium.
 FT ACT SITE 155 By similarity.
 FT METAL 242 Magnesium (By similarity).
 FT METAL 243 Magnesium (By similarity).
 FT METAL 283 Magnesium (By similarity).
 FT METAL 310 Magnesium (By similarity).
 SQ SEQUENCE 426 AA; 45483 MW; 63AED7B7A4E15001 CRC64;

Query Match 74.0%; Score 1595.5; DB 1; Length 426;
 Best Local Similarity 74.5%; Pred. No. 1.4e-86;
 Matches 313; Conservative 47; Mismatches 59; Indels 1; Gaps 1;

QY 7 VFAEILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLKGVLYK 65
 DB 7 VFAEILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYLKGVLYK 66
 QY 66 AVENNEEIGDELADDDORLIDEAMIKLDGTANKSRIGANAILGVMAVAKAADSA 125
 DB 67 AVLAVIEQIGELVGYDATEQRLDQAMFDDLDATDNKSLGANAILGVSLVAHAASEAS 126
 QY 126 GLPLFRYIGGNPAHVLVPMNNIINGGAHDSGVDDVDFEAPITGAETFESEALRNGAEVY 185
 DB 127 DLPLFRYIGGNPAHLLVPMNNIINGGSHADSNVDIQEFMIAPICASFSEALRWGAEVY 186
 QY 186 HALKSIVKEKSLTGLDGEFGFAPSVGSTREALDLIVKAIKAGFTPGKDIALDALDVASS 245
 DB 187 HTLKKVLKNKGLATGLDGEFGFAPNLGSNREALDLILEAKGYTPGEQIALDALDVASS 246
 QY 246 EFFKDGTYHFGGQHSAAEMANVVAELVDAYPIVSIEDPLQEDDWEQYTNLTATIGDKVQ 305
 DB 247 EFKDGTYHFGGQHSAAEMTEYVAELVAYPLVSIEDPLQEDDWEQYTNLTATIGDKVQ 306
 QY 306 IVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGVTSMMHSRG 365
 DB 307 LVGDDLVFTNPERLARGIEENSANALLVKVNIQIGSLTETLDVAVELAQNGFKCMHSRG 366
 QY 366 ETEDTTIADLAVNLCQIKTGAPARSDRVAKYNQLLRIBQLLDAGVYAGRSAPFRFQ 425
 DB 367 ETEDVTIADLAVNLCQIKTGAPARSDRVAKYNQLLRIBQLLDAGVYAGRSAPFRFQ 426

RESULT 5
 Q6ADR6 PRELIMINARY; PRT; 426 AA.
 AC Q6ADR6
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Enolase.
 GN Name=eno; OrderedLocusNames=Lxxl7200;
 OS Leifsonia xyl (subsp. xyl).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Microbacteriaceae; Leifsonia.
 OX NCBI_TaxID=59736;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CTCB07;
 RX PubMed=15305603;
 RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
 RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
 RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
 RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
 RA Almeida N.F. Jr., Carver H., Coutinho L.L., El-Dorry H.A.,
 RA Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
 RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,
 RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
 RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
 RA Tsai S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.,
 RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
 RL xyl subsp. xyl.",
 RL Mol. Plant Microbe Interact. 17:827-836(2004).
 CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 CC H(2O).
 CC -I- COFACTOR: Magnesium is required for catalysis and for stabilizing

CC the dimer (By similarity).
 CC -I- PATHWAY: Glycolysis.
 CC -I- SUBUNIT: Homodimer (By similarity).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -I- SIMILARITY: Belongs to the enolase family.
 DR EMBL: AE016822; AAT89480.1; -;
 DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR000941; Enolase.
 DR Pfam: PF01113; Enolase_C; 1.
 DR PRINTS: PR00148; ENOLASE.
 DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 426 AA; 45227 MW; 43DE1F7B691BC2E1 CRC64;

Query Match 72.7%; Score 1567.5; DB 2; Length 426;
 Best Local Similarity 72.8%; Pred. No. 6.3e-85;
 Matches 310; Conservative 47; Mismatches 68; Indels 1; Gaps 1;

QY 1 VASIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYL 59
 DB 1 MAEIAVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRDGG-DRYL 60
 QY 60 GKGVKAVENVNBEIGDELADDDORLIDEAMIKLDGTANKSRIGANAILGVMAVAK 119
 DB 61 GKGVKAVDAVLDEIGPAIEGFEASDQRIVDAMIELDGTDNKKRIGANAILGVSLAVAK 120
 QY 120 AADSSAGLPLFRYIGGNPAHVLVPMNNIINGGAHDSGVDDVDFEAPITGAETFESEALR 179
 DB 121 AADSSADLPLFRYIGGNPAHVLVPMNNIINGGAHDSGVDDVDFEAPITGAETFESEALR 180
 QY 180 NGAEVYHALKSIVKEKSLTGLDGEFGFAPSVGSTREALDLIVKAIKAGFTPGKDIALA 239
 DB 181 WGVETVHSLKALLKSKLNTGLDGEFGFAPLEHNRALDLTAEAEKAGFTVGSQIALG 240
 QY 240 LVASSEFPKDGTYHFGGQHSAAEMANVVAELVDAYPIVSIEDPLQEDDWEQYTNLTAT 299
 DB 241 LVVASTFEFFGVYRFGQDRTAEMSAYYTLANNYPLVSIEDPLQEDDWEQYTNLTAT 300
 QY 300 IGDQVQIVGDDPFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVDMHAGVTSMM 359
 DB 301 IGSLQLVGGDDLFTNPERLKEGIAKKAANSILVKVNIQIGTLTETLDVAVELAQNGFKCM 360
 QY 360 MSHRSGEETDTIADLAVNLCQIKTGAPARSDRVAKYNQLLRIBQLLDAGVYAGRS 419
 DB 361 LSHRSGEETDTIADLAVNLCQIKTGAPARSDRVAKYNQLLRIBQLLDAGVYAGRS 420
 QY 420 FPRFQ 425
 DB 421 FPRYQ 426

RESULT 6
 Q741U7 PRELIMINARY; PRT; 429 AA.
 ID Q741U7
 AC Q741U7
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Eno.
 GN Name=eno; OrderedLocusNames=MAP0990;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=k10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.,

```

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC
CC -|- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -|- PATHWAY: Glycolysis.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the enolase family.
DR EMBL: AE017230; AA033107.1; -.
DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:0046344; P:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; F:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF001113; Enolase_N; 1.
DR Pfam: PF03952; Enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 429 AA; 44873 MW; C5B7C0460F01D703 CRC64;

Query Match 72.4%; Score 1560.5; DB 2; Length 429;
Best Local Similarity 72.9%; Pred. No. 1.6e-84;
Matches 307; Conservative 47; Mismatches 66; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCASTGVHEAHELRDGGRYLGKV 63
DB 4 IEQVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCASTGVHEAHELRDGGRYLGKV 63

QY 64 LKAVERNBERIGDELAGEADDORLIDEAMIKLDGTANKSRLGANALGVSMVAKAAAD 123
DB 64 QKAVQAVLDEIGPANGVGLNADDQRLVDQALVDLDTGTPKSRGLGNALGVSLAVAKAAD 123

QY 124 SAGLPFRYTGNNVHVLVPMNMINGGAHSGVDVQVFMIAPIGAEFTSEALRWAE 183
DB 124 SABLPLFRYLGPNVHILVPMNMILNGGAHADTAVDIQEFMVAPIGAFSAELRWAE 183

QY 184 VYHALKSIVKEGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGTPKDTALALDVA 243
DB 184 VYHSLSVLKKGLSTGLGDEGGFADPVAGTAAALDLIGRAIESAGFKLGTVALALDAA 243

QY 244 SSEFFKDGDT-YHPEGGHSAEAMVAVYAEIVDAPYIVSIEDPLQEDDWEQYTNLTATGD 302
DB 244 ATFFSDGTGYKEGSTRTAQWAEFYAGLLGAYPLVSIEDPLSEDDWDGWAALTASIGD 303

QY 303 KVOIVGDDFFVTNPERLKEGIKAAKANSILVKVQNQIGTLTETFDVDMHAGRYTSMWSH 362
DB 304 RVQLVGDDVFTNPERLEEGIEKGVANALLVKVQNQIGTLTETLDAVALAHSGVRYTSMWSH 363

QY 363 RSGETEDTTIADLAVLNCCQIKTGAPARSDRVAKYNQLRIEQLLDGAGVYAGRSAPFR 422
DB 364 RSGETEDTTIADLAVAGVGQIKTGAPARSDRVAKYNQLRIEALGDAARYAGDLAPFR 423

QY 423 F 423
DB 424 F 424

RESULT 7
ID ENO_MYCTU STANDARD; PRT; 429 AA.
AC P96377;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocName=Rv1023, MT1051; ORFName=MTCY10G2.26c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

```

```

OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A., R.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Stulsen J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwin M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -|- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -|- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -|- PATHWAY: Glycolysis.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the enolase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: BX842575; CAB06856.1; -.
CC EMBL: AE000516; AAK45302.1; -.
CC PIR: B70623; B70623.
CC HSSP: O9NDH8; 10EP.
CC TIGR: MT1051; -.
CC Tuberculist: Rv1023; -.
CC HAMAP: MF_00318; -.
CC InterPro: IPR000941; Enolase.
CC Pfam: PF00113; Enolase_C; 1.
CC Pfam: PF03952; Enolase_N; 1.
CC PRINTS: PR00148; ENOLASE.
CC ProDom: PD000902; Enolase; 1.
CC TIGRFAMs: TIGR01060; eno; 1.
CC PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
FT ACT_SITE 154 154 By similarity.
FT METAL 241 241 Magnesium (By similarity).
FT METAL 283 283 Magnesium (By similarity).
FT METAL 310 310 Magnesium (By similarity).
SQ SEQUENCE 429 AA; 44961 MW; 5AC59531AC9F00E8 CRC64;

Query Match 72.1%; Score 1553.5; DB 1; Length 429;
Best Local Similarity 73.2%; Pred. No. 4.3e-84;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSCASTGVHEAHELRDGGRYLGKV 63

```

```

Db      4 IEQVAREILDSRGNTVEVEVALIDGTTFARAAVPSGASTGBHEAVELRDGGDRYGGKGV 63
QY      64 LKAVERNNEIIGDELGLAGLEADQRLIDEAMIKLDGTANKSRILGANILGVSMVAKAAAD 123
Db      64 QKAVQAVLDEIGPAVIGLGNADQRLVDQALVDLDTGTPDKSRILGNAILGVSLAVAKAAAD 123
QY      124 SAGLPLFRYIGGPNNAHLVPPMMNIINGGAHDSGVVDVQEFMIPIGAEFTFSEALRWGAE 183
Db      124 SAGLPLFRYIGGPNNAHLVPPMMNIINGGAHADTAVDIQEFMVAPIGAPSFVEALRWGAE 183
QY      184 VYHALKSVLKEKGLSTGLGDEGFAPSVGSTREALDLIVKATEKAGFTPGKDIALALDVA 243
Db      184 VYHALKSVLKEKGLSTGLGDEGFAPDVAGTTAAALDLISRAIESAGLRPGADVALALDAA 243
QY      244 SSEFFKDG-T-YHPEGGQHSAAEMANVYAEIVDAPYIVSTEDPLQDDWEGYTNLTATIGD 302
Db      244 ATEFFTDGTGYVFEFTTADQMTQEFYAGLLGAYPLVSTEDPLSDDDWGWAAALTASIGD 303
QY      303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHAGYTSMMSH 362
Db      304 RVOIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETLDVTLAHHGGYRTMISH 363
QY      363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAPPR 422
Db      364 RSGETEDTTIADLAVAIAGSGQIKTGAPARSERVAKYNQLLRIEALGDAAARYAGDLAFPR 423
QY      423 F 423
Db      424 F 424

```

RESULT 8

```

Q7U0U6
ID Q7U0U6 PRELIMINARY; PRT; 429 AA.
AC Q7U0U6;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE PROBABLE ENOLASE ENO (EC 4.2.1.11).
GN Name=eno; OrderedLocusNames=Mb1051;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Actin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis."
RA Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the enolase family.
DR EMBL: BX248337; CAD93911.1; -.
DR HSP: P08324; I291.
DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO: GO:0016829; F:lyase activity; IEA.
DR GO: GO:0000287; F:magnesium ion binding; IEA.
DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; Enolase_C; 1.
DR Pfam: PF03952; Enolase_N; 1.

```

```

DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 429 AA; 44862 MW; F808FB8E0D891C9A CRC64;

```

```

Query Match 72.1%; Score 1553.5; DB 2; Length 429;
Best Local Similarity 73.2%; Pred. No. 4.3e-84;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

```

```

QY      4 IMHVAREILDSRGNTVEVEVALIDGTTFARAAVPSGASTGBHEAVELRDGGDRYGGKGV 63
Db      4 IEQVAREILDSRGNTVEVEVALIDGTTFARAAVPSGASTGBHEAVELRDGGDRYGGKGV 63
QY      64 LKAVERNNEIIGDELGLAGLEADQRLIDEAMIKLDGTANKSRILGANILGVSMVAKAAAD 123
Db      64 QKAVQAVLDEIGPAVIGLGNADQRLVDQALVDLDTGTPDKSRILGNAILGVSLAVAKAAAD 123
QY      124 SAGLPLFRYIGGPNNAHLVPPMMNIINGGAHDSGVVDVQEFMIPIGAEFTFSEALRWGAE 183
Db      124 SAGLPLFRYIGGPNNAHLVPPMMNIINGGAHADTAVDIQEFMVAPIGAPSFVEALRWGAE 183
QY      184 VYHALKSVLKEKGLSTGLGDEGFAPSVGSTREALDLIVKATEKAGFTPGKDIALALDVA 243
Db      184 VYHALKSVLKEKGLSTGLGDEGFAPDVAGTTAAALDLISRAIESAGLRPGADVALALDAA 243
QY      244 SSEFFKDG-T-YHPEGGQHSAAEMANVYAEIVDAPYIVSTEDPLQDDWEGYTNLTATIGD 302
Db      244 ATEFFTDGTGYVFEFTTADQMTQEFYAGLLGAYPLVSTEDPLSDDDWGWAAALTASIGD 303
QY      303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVDMHAGYTSMMSH 362
Db      304 RVOIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETLDVTLAHHGGYRTMISH 363
QY      363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLDGAGVYAGRSAPPR 422
Db      364 RSGETEDTTIADLAVAIAGSGQIKTGAPARSERVAKYNQLLRIEALGDAAARYAGDLAFPR 423
QY      423 F 423
Db      424 F 424

```

RESULT 9

```

Q82HH5
ID Q82HH5 PRELIMINARY; PRT; 428 AA.
AC Q82HH5;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Putative enolase.
GN Name=eno; OrderedLocusNames=SAV3533;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RA avermitilis: deducing the ability of producing secondary
RA metabolites."
RA Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;

```

"Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*,"
 Nat. Biotechnol. 21:526-531(2003).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H₂O.
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 DR ENBL; AP005035; BAC71245.1; -.
 DR HSP; Q9NDH8; IOEP.
 DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 DR GO: GO:000287; F:magnesium ion binding; IEA.
 DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF03952; Enolase_C; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR TIGRFAMs; TIGR01060; eno; 1.
 DR PROSITE; PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 428 AA; 45869 MW; 095A3976F6C5F70D CRC64;
 Query Match 71.9%; Score 1550.5; DB 2; Length 428;
 Best Local Similarity 73.5%; Pred. No. 6.4e-84;
 Matches 308; Conservative 43; Mismatches 67; Indels 1; Gaps 1;
 QY 7 VFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGG-DRYLGGKVLK 65
 DB 7 VFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGG-DRYLGGKVLK 66
 QY 66 AVENVNEEIGDELGLAGLADQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAADS 125
 DB 67 AVLAVIEQIGPELVGYDTEQRLIDQAMFDDATDNKSLGANAILGVSLAVAHAAEAS 126
 QY 126 GLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALRNGAEVY 185
 DB 127 DLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALRNGAEVY 186
 QY 186 HALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALADVASS 245
 DB 187 HTLUKKVLKTLGSLGDEGGFAPNLESNRALDLIIKAIKAGFTPGKDIALADVASS 246
 QY 246 EFPKDGTYHFEFGGSHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATIGDKVQ 305
 DB 247 EFKDGKYEFGKRSAAEMTEYELVAYPLVSIEDPLQEDDWEQYTNLTATIGDKVQ 306
 QY 306 IVGDDPFVTPNPERKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSNMHRSRG 365
 DB 307 IVGDDPFVTPNPERKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSNMHRSRG 366
 QY 366 ETEDTTIADLAVNLCCQIKTGAPARSDRVAKYNOLLRIEQLICDAGVYAGRAFFRFQ 424
 DB 367 ETEDVTIADLAVAVNCQIKTGAPARSDRVAKYNOLLRIEQLICDAGVYAGRAFFRFQ 425
 RESULT 10
 Q6AAB8
 ID Q6AAB8 PRELIMINARY; PRT; 426 AA.
 AC Q6AAB8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Enolase (EC 4.2.1.11)
 GN OrderedLocusNames=PPA0545;
 OS Propionibacterium acnes.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Propionibacterineae; Propionibacteriaceae; Propionibacterium.

NCBI_TaxID=1747;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=KPA171202 / DSM 16379;
 RX PubMed=15286373; DOI=10.1126/science.1100330;
 RA Brueggemann H., Henne A., Hofer F., Liesegang H., Wierse A.,
 RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.,
 RT "The complete genome sequence of *Propionibacterium acnes*, a commensal of human skin,"
 RL Science 305:671-673(2004).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate + H₂O.
 CC -1- COFACTOR: Magnesium is required for catalysis and for stabilizing the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 DR ENBL; AE017283; AAT82298.1; -.
 DR GO: GO:0000015; C:phosphopyruvate hydratase complex; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 DR GO: GO:0004634; F:phosphopyruvate hydratase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF03952; Enolase_C; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR TIGRFAMs; TIGR01060; eno; 1.
 DR PROSITE; PS00164; ENOLASE; 1.
 KW Complete proteome; Glycolysis; Lyase; Magnesium.
 SQ SEQUENCE 426 AA; 45530 MW; 39FD0E9A2941123F CRC64;
 Query Match 69.9%; Score 1505.5; DB 2; Length 426;
 Best Local Similarity 68.9%; Pred. No. 2.9e-81;
 Matches 293; Conservative 58; Mismatches 73; Indels 1; Gaps 1;
 QY 1 VAEIMVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGG-DRYL 59
 DB 1 MATIEFIEAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGG-DRYL 60
 QY 60 GKGLKAVENNVNEEIGDELGLAGLADQRLIDEAMIKLDGTANKSRILGANAILGVSMVAK 119
 DB 61 GKGLKAVENNVNEEIGDELGLAGLADQRLIDEAMIKLDGTANKSRILGANAILGVSMVAK 120
 QY 120 AAADSLGLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALR 179
 DB 121 AAADSLGLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALR 180
 QY 180 NGAEVYHALKSVIKKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALA 239
 DB 181 WGAATVTHSLKLVKLDKGLATGLGDEGGFAPNLESNRALDLIIKAIKAGFTPGKDIALA 240
 QY 240 LDVASSEFFKDGTYHFEFGGSHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTAT 299
 DB 241 LDVASSEFFKDGTYHFEFGGSHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTAT 300
 QY 300 IGDKVQIVGDDPFVTPNPERKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSM 359
 DB 301 LGSEKIQIVGDDPFVTPNPERKEGIAKKAANSILVKVNIQIGTLTETFDVDMHARAGYTSM 360
 QY 360 MSHRSGEETDTTIADLAVNLCCQIKTGAPARSDRVAKYNOLLRIEQLICDAGVYAGRA 419
 DB 361 MSHRSGEETDTTIADLAVNLCCQIKTGAPARSDRVAKYNOLLRIEQLICDAGVYAGRA 420
 QY 420 FPRFQ 424
 DB 421 FPRFQ 425
 RESULT 11
 ENO_MYCLE
 ID ENO_MYCLE STANDARD; PRT; 429 AA.

Q9CD42;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
 glycerate hydro-lyase).
 GN Name=eno; OrderedLocusNames=ML0255;
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
 RA Cole S.T., Eiglmier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
 RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 H(2)O.
 CC -1- COPACTOR: Magnesium is required for catalysis and for stabilizing
 the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL583917; CAC29763.1; ALT_INIT.
 CC HSP; Q9NDH8; 10EP.
 CC Leprona; ML0255; -.
 CC HAMAP; MF_00318; -; 1.
 CC InterPro; IPR000941; Enolase.
 CC Pfam; PF00113; Enolase_C; 1.
 CC Pfam; PF03932; Enolase_N; 1.
 CC PRINTS; PR00148; ENOLASE.
 CC ProDom; PD000902; Enolase; 1.
 CC TIGRFAMs; TIGR01060; eno; 1.
 CC PROSITE; PS00164; ENOLASE; 1.
 CC Complete proteome; Glycolysis; Lyase; Magnesium.
 CC ACT_SITE 154 By similarity.
 CC METAL 241 Magnesium (By similarity).
 CC METAL 283 Magnesium (By similarity).
 CC METAL 310 Magnesium (By similarity).
 CC SEQUENCE 429 AA; 45201 MW; A8299C1EA44617B5 CRC64;
 Query Match 59.7%; Score 1501.5; DB 1; Length 429;
 Best Local Similarity 71.0%; Pred. No. 5.1e-81;
 Matches 299; Conservative 47; Mismatches 74; Indels 1; Gaps 1;
 QY 4 IMHVFAREILDSRGNTVEAEVFLDGDGSHGACVPSGAGTGVHEAEHLRDGGRVYLGKV 63
 Db 4 IEQVGAREILDSRGNTVEAEVFLDGTFAAAVPSGAGTGVHEAEHLRDGGRVYLGKV 63
 QY 64 LKAVERNNEIGDELGLAGLADDQRLIDEAMIKLDGTANKSRGANAILGVSAVAKAAAD 123
 Db 64 KRAVDVAVLDEIGPVGWVGLNANDQRLIDQGLDGTGTPKSRGNAILGVSLAVAKAAAD 123

124 SAGLPLFRYGGPNNAHVLVPVPMNIINGGAHAGSDVDQVEFMIAPIGAETFESEALRNGAE 183
 124 SASELPLFRYGGNAHILPVPVPMNIIINGGAHADTAVDQVEFMVAPIGAPSFVEALRWGAE 183
 184 VYHALKSIVKEKGLSTGLGDEGFAPSVGSTRALDLIVKAIKAGTTPGKDIALADVA 243
 184 VYHALKSIVLKKGLSTGLGDEGFAPEVAGTTAALDLVTLAIEAAGFKPGADVALALDAA 243
 244 SBEFFKDGTT-YHFEGGGSHAAEMANVYAEIVDSIEDPLQEDDDEGTYNLTATIGD 302
 244 ATEFYTDGIGYHFEFGMTADQWTEFVADLLGSGYPLVSTEDPLSEDDWDGWAALTASIGE 303
 303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQITLTETETPDVDMHARAGYTMMSH 362
 304 QVOIVGDDIFATNPERLEEGIRGVANALLVKVNIQITLTETLEVALAHSGYRTMISH 363
 363 RSETEDTTTADLAVALNCQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRAFPR 422
 364 RSETEDTMIADLVVALGSGQIKTGAPARSERVAKYNOLLRIEELGDAAARYAGDLAFLR 423
 423 F 423
 424 Y 424

RESULT 12
 Q83H73
 ID Q83H73 PRELIMINARY; PRT; 428 AA.
 AC Q83H73;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Enolase (EC 4.2.1.11).
 GN Name=eno; OrderedLocusNames=TW793;
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcineae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=218496;
 RN [1]_TaxID=218496;
 RP SEQUENCE FROM N.A.
 RC STRAIN=TW08/27;
 RX MEDLINE=22495039; PubMed=12606174; DOI=10.1016/S0140-6736(03)12597-4;
 RA Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
 RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,
 RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
 RA Barrell B.G., Parkhill J., Rellman D.A.;
 RT "Sequencing and analysis of the genome of the Whipple's disease
 bacterium Tropheryma whipplei.";
 RL Lancet 361:637-644(2003).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 H(2)O.
 CC -1- COPACTOR: Magnesium is required for catalysis and for stabilizing
 the dimer (By similarity).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the enolase family.
 DR EMBL; BX251412; CAD67452.1; -.
 DR HSP; Q9NDH8; 10EP.
 DR GO; GO:0000015; C:phosphoryruvate hydratase complex; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0000287; F:magnesium ion binding; IEA.
 DR GO; GO:0004634; F:phosphoryruvate hydratase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000941; Enolase.
 DR Pfam; PF00113; Enolase_C; 1.
 DR Pfam; PF03952; Enolase_N; 1.
 DR PRINTS; PR00148; ENOLASE.
 DR ProDom; PD000902; Enolase; 1.
 DR TIGRFAMs; TIGR01060; eno; 1.
 DR PROSITE; PS00164; ENOLASE; 1.
 DR Complete proteome; Glycolysis; Lyase; Magnesium.
 KW COMPLETE PROTEOME; Glycolysis; Lyase; Magnesium.
 SEQUENCE 428 AA; 45961 MW; 356D63DCCC24B581 CRC64;

```
Query Match 66.6%; Score 1436; DB 2; Length 428;
Best Local Similarity 67.9%; Pred. No. 3.9e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

QY 7 VFAREILDSRGNTVEAEVFLDGGSHGVAGVPSGASTGVHAEHLRD-GGDYILGKGVLK 65
DB 7 VHARQILDSRGTPTEVEVTLDDGSMGRSAVPSGASTGAFEAHLRDQDNNEYLKGVTR 66
QY 66 AVENVNEEIGDELAGLEADQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAADSA 125
DB 67 AVRSVNSEIAPVLIGFADFQVGLDHRMIELDGTNNKSRILGANAILGVSLASAAARAA 126
QY 126 GLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALRNGAEVY 185
DB 127 DLSLFRYLGGPSSRILVPMNMNIINGGAHDTGVDIQEFMILPVGARSFSESRLWGCEVY 186
QY 186 HALKSUIKEKLGSTGLDGGFPAPSVGSTREALDLIVKAIEKAGFTPGKDIALLDVASS 245
DB 187 HSLKVQLRESGLSGLDGGFPALRSNRALTDLILSAIEKAGFSPGIDIVLALDIAAS 246
QY 246 EFK-DGTYHFEFGQHSAAEMANVYAEVDAYPTVSTEDPLODDWEGYTNLTATIGDKV 304
DB 247 EFCAPGHYRFEKGDTSDLSIYEGLLSSYPLVSTEDPLODDWEGYRTLTTHIGDRV 306
QY 305 QIVGDDFFVNTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAAGYTSMMSHRS 364
DB 307 QIVGDDLVFVNTSRLSRGIQSGVANSILVKVNOIGTLTETFDVADMAKGGYTAVLSHRS 366
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRLEQLLGDAGVYAGRSAPFRFQ 424
DB 367 GETEDTTIADMAVATNCGQIKTGAPARGERIAKYNQLLRLEEKLRGARSARYAGWLSYPRWQ 426
QY 425 G 425
DB 427 G 427

RESULT 13
ID Q83FF7 PRELIMINARY; PRT; 446 AA.
AC Q83FF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Enolase (EC 4.2.1.11).
GN Name=eno; OrderedLocuNames=TW7783;
OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
OX NCBI_TaxID=203267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Twist;
RX MEDLINE=22784088; PubMed=12902375;
RA Reult D., Ogata H., Audic S., Robert C., Suhre K., Drancourt M.,
RA Claverie J.-M.
RT "Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a
RT reduced genome."
RL Genome Res. 13:1800-1809(2003).
CC -I- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -I- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC -I- PATHWAY: Glycolysis.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the enolase family.
DR EMBL; AE016852; AA044880.1; --
DR HSP; Q9NDH8; 10EP.
DR GO; GO:0000015; C:phosphopyruvate hydratase complex; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.

DR GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; Enolase_C; 1.
DR Pfam; PF03952; Enolase_N; 1.
DR PRINTS; PR00148; ENOLASE.
DR PRODOM; PD000902; Enolase; 1.
DR TIGRFAMs; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Complete proteome; Glycolysis; Lyase; Magnesium.
SQ SEQUENCE 446 AA; 48141 MW; 786E73F99F94A6DA CRC64;

Query Match 66.6%; Score 1436; DB 2; Length 446;
Best Local Similarity 67.9%; Pred. No. 4.1e-77;
Matches 286; Conservative 51; Mismatches 82; Indels 2; Gaps 2;

QY 7 VFAREILDSRGNTVEAEVFLDGGSHGVAGVPSGASTGVHAEHLRD-GGDYILGKGVLK 65
DB 25 VHARQILDSRGTPTEVEVTLDDGSMGRSAVPSGASTGAFEAHLRDQDNNEYLKGVTR 84
QY 66 AVENVNEEIGDELAGLEADQRLIDEAMIKLDGTANKSRILGANAILGVSMVAKAAADSA 125
DB 85 AVRSVNSEIAPVLIGFADFQVGLDHRMIELDGTNNKSRILGANAILGVSLASAAARAA 144
QY 126 GLPLFRYIGGNPAHVLVPMNMNIINGGAHDSGVVDVQEFMIAPIGAEFTFSEALRNGAEVY 185
DB 145 DLSLFRYLGGPSSRILVPMNMNIINGGAHDTGVDIQEFMILPVGARSFSESRLWGCEVY 204
QY 186 HALKSUIKEKLGSTGLDGGFPAPSVGSTREALDLIVKAIEKAGFTPGKDIALLDVASS 245
DB 205 HSLKVQLRESGLSGLDGGFPALRSNRALTDLILSAIEKAGFSPGIDIVLALDIAAS 264
QY 246 EFK-DGTYHFEFGQHSAAEMANVYAEVDAYPTVSTEDPLODDWEGYTNLTATIGDKV 304
DB 265 EFCAPGHYRFEKGDTSDLSIYEGLLSSYPLVSTEDPLODDWEGYRTLTTHIGDRV 324
QY 305 QIVGDDFFVNTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAAGYTSMMSHRS 364
DB 325 QIVGDDLVFVNTSRLSRGIQSGVANSILVKVNOIGTLTETFDVADMAKGGYTAVLSHRS 384
QY 365 GETEDTTIADLAVALNCQIKTGAPARSDRVAKYNQLLRLEQLLGDAGVYAGRSAPFRFQ 424
DB 385 GETEDTTIADMAVATNCGQIKTGAPARGERIAKYNQLLRLEEKLRGARSARYAGWLSYPRWQ 444
QY 425 G 425
DB 445 G 445

RESULT 14
ENO_AGRT5
ID ENO_AGRT5 STANDARD; PRT; 424 AA.
AC Q8UFH1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN Name=eno; OrderedLocuNames=Atu1426, AGR_C_2631;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
```



```
FT METAL 242 242 Magnesium (By similarity).
FT METAL 285 285 Magnesium (By similarity).
FT METAL 312 312 Magnesium (By similarity).
SQ SEQUENCE 426 AA; 45245 MW; 6F2E0B4816A22458 CRC64;

Query Match 65.5%; Score 1412.5; DB 1; Length 426;
Best Local Similarity 66.4%; Pred. No. 9.5e-76;
Matches 280; Conservative 59; Mismatches 80; Indels 3; Gaps 2;

Qy 1 VAEIMHVFAREILDSRGNPTVEARVFLDDGSHGVAGVPSGASTGVHEAHELROGGD-RYL 59
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1 MTEIVDIIAREILDSRGNPTVEVDVLEDAFGRAAVPSGASTGAHEANEKROGDKARYL 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 60 KGVLKAVENVNEEIGDELAGEADDQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAVAK 119
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 61 KGGVQQAADAVNGEIPDALSGVDAEDQRRVDNLMIELDGTNPKNKARLGANAILGVSLATAK 120
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 120 AAADSAGLPLFRYIGGNVHVLVPMNMNIINGNAHDSGVVDVQEFPMIADIGAEFFSEALR 179
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 121 AAASAGLPLYKYVGVNARVLPMPMNIINGGAHADNPDIQEFMLPTGAKDFREGLR 180
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 180 NGAEVVHALKSVIKEKGLSTGLDGGGFPSPVSGSTREALDLIVKAIKAGFTPGKDIATA 239
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 181 MGAEIFHALKALKADAGHNTNVGDEGGFAPNLASAEAAALDFIVKAGEKAGTKAGDDFVLG 240
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 240 LDVASSEFFKDGTYHFEGQHS--AAEMANYAELVDAYPIVSIEDPLQEDDWEGYTNLT 297
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 241 LDVASTEFFKNGKYELESGKSLDPAAMVDYLAGLVAKFPILITIEDGMAEDDFDGWKLTT 300
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 298 ATGDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKNOIGLITETTFDAVDMAHRAGYT 357
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 301 DTLGKKVQLVGDDLFVTNPKRLQMLDKGLANSILVKVNQIGTILSETIDAVELAHRHGYT 360
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 358 SMMSHRSGETETDIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGAGVYAGR 417
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 361 SVMSHRSGETETDIADLAVALNCGQIKTGLSKASDRDTAKYNOLLRIEQLLDGQVYAGR 420
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 418 SA 419
Db :|
Qy 421 AA 422
```

Search completed: September 29, 2005, 12:13:47
Job time : 89.5 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 09:39:18 ; Search time 38.5 Seconds
(without alignments)
824.048 Million cell updates/sec

Title: US-10-728-947-4
Perfect score: 2155
Sequence: 1 VAEIMHVPFAREILDSRGNT.....QLLGAGVYGRSAFFRFG 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/6C COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	US-09-860-768-4	Sequence 4, Appli
2	2151	99.8	425	US-09-860-768-2	Sequence 2, Appli
3	1357.5	63.0	432	US-09-302-540-12060	Sequence 12060, A
4	1306.5	60.6	434	US-09-710-279-2222	Sequence 2222, Ap
5	1306.5	60.6	440	US-09-134-001C-5641	Sequence 5641, Ap
6	1280.5	59.4	434	US-09-830-217-22	Sequence 22, Appl
7	1280.5	59.4	434	US-10-278-946-22	Sequence 22, Appl
8	1278.5	59.3	444	US-09-252-991A-21020	Sequence 21020, A
9	1256.5	58.3	434	US-09-583-110-3817	Sequence 3817, Ap
10	1240	57.5	398	US-09-107-532A-3810	Sequence 3810, Ap
11	1231.5	57.1	444	US-09-328-352-7269	Sequence 7269, Ap
12	1226.5	56.9	420	US-08-847-065-19	Sequence 19, Appl
13	1226.5	56.9	420	US-09-829-382-19	Sequence 19, Appl
14	1222	56.7	459	US-09-489-039A-10061	Sequence 10061, A
15	1201.5	55.8	445	US-09-540-236-2354	Sequence 2354, A
16	1186.5	55.1	456	US-09-543-681A-6090	Sequence 6090, Ap
17	1168.5	54.2	434	US-09-438-185A-802	Sequence 802, App
18	1131.5	52.5	407	US-09-107-433-4986	Sequence 4986, Ap
19	1097	50.9	443	US-09-949-016-8359	Sequence 8359, Ap
20	1095	50.8	434	US-09-949-016-6153	Sequence 6153, Ap
21	942.5	43.7	452	US-09-198-452A-853	Sequence 853, App
22	902	41.9	311	US-08-961-083-184	Sequence 184, App
23	902	41.9	301	US-09-536-784-184	Sequence 184, App
24	651.5	30.2	301	US-09-248-796A-17053	Sequence 17053, A
25	504.5	23.4	177	US-08-446-920-11	Sequence 11, Appl
26	458	21.3	148	US-09-248-796A-17054	Sequence 17054, A
27	265	12.3	113	US-09-513-999C-5420	Sequence 5420, Ap

ALIGNMENTS

RESULT 1

US-09-860-768-4
; Sequence 4, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-4

Query Match	100.0%;	Score 2155;	DB 4;	Length 425;
Best Local Similarity	100.0%;	Pred. No. 1.8e-190;		
Matches 425;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VAEIMHVPFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGDRLG	60	
Db	1	VAEIMHVPFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDDGGDRLG	60	
Qy	61	KGVLKAVENVNBEIGDELADQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA	120	
Db	61	KGVLKAVENVNBEIGDELADQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA	120	
Qy	121	AADSAGLPLFRYIGGPNHVLVPMNMNINGGAHDSGVVDQEFMFIAPIGAEFTFSALRN	180	
Db	121	AADSAGLPLFRYIGGPNHVLVPMNMNINGGAHDSGVVDQEFMFIAPIGAEFTFSALRN	180	
Qy	181	GAEVYHALKSVTKKGLSTGLDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL	240	
Db	181	GAEVYHALKSVTKKGLSTGLDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL	240	
Qy	241	DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTATI	300	
Db	241	DVASSEFFKDGTYHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPQEDDWEGYTNLTATI	300	
Qy	301	GDVKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETTFDAVDMAHRAGYTSM	360	
Db	301	GDVKQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETTFDAVDMAHRAGYTSM	360	


```

QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
QY 421 PRFQ 425
Db 421 PRFQ 425

RESULT 2
US-09-860-768-2
; Sequence 2, Application US/09860768
; Patent No. 6713289
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleoside Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-2

```

```

Query Match 99.8%; Score 2151; DB 4; Length 425;
Best Local Similarity 99.8%; Pred. No. 4.3e-190;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGGDRYL 60
Db 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGGDRYL 60
QY 61 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAKA 120
QY 121 AADSAGLPLFRYIGGPNNAHVLPPVPMNIIINGGAHADSGVDVQEFMTAPIGAETTFSEALRN 180
Db 121 AADSAGLPLFRYIGGPNNAHVLPPVPMNIIINGGAHADSGVDVQEFMTAPIGAETTFSEALRN 180
QY 181 GAEVYHALKSVIKEKGLSTGLGDEGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Db 181 GAEVYHALKSVIKEKGLSTGLGDEGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
QY 241 DVASSEFFKDGTVHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDMEGYTNLTATI 300
Db 241 DVASSEFFKDGTVHFEGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDMEGYTNLTATI 300
QY 301 GDKVQIVGGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMHRAAGTSMM 360
Db 301 GDKVQIVGGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMHRAAGTSMM 360
QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF 420
QY 421 PRFQ 425
Db 421 PRFQ 425

```

```

RESULT 3
US-09-902-540-12060
; Sequence 12060, Application US/09902540

```

```

; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12060
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12060

Query Match 63.0%; Score 1357.5; DB 4; Length 432;
Best Local Similarity 63.8%; Pred. No. 8.4e-117;
Matches 271; Conservative 59; Mismatches 90; Indels 5; Gaps 2;

QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG-DRYL 59
Db 1 MTEISQILAREVLDSRGNTVEAEVQLAGGARAAVPSGASTGEHEAIELRDGDKHRYL 60
QY 60 KGVLKAVENNVNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRKLGANAILGVSMVAK 119
Db 61 KGVLKAVENNVVDLAPALVGMDDAADOFAVDQRMLELDGTATKGLGANAILAVSMAAAR 120
QY 120 AADSAGLPLFRYIGGPNNAHVLPPVPMNIIINGGAHADSGVDVQEFMTAPIGAETTFSEALR 179
Db 121 AADAAGHGLPLYRYGVGVQARTLPVPLMNTLNGGAHADTRVDVQEFMVVPAGASSFAEGLR 180
QY 180 NGAEVYHALKSVIKEKGLSTGLGDEGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALA 239
Db 181 WGAEVYHALKSVIKEKGLSTGLGDEGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALA 240
QY 240 LDVASSEFFKDGTVHF-EGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDMEGYTN 295
Db 241 LDVASSEFFKDGSKYKLGEGKEYDSTGLLEYRYGLSERYPIISIEDGMAEDDMEGWKK 300
QY 296 LTATTGDKVQIVGGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETTFDAVDMHRAAG 355
Db 301 LTDALGSRMQLVGGDDLFTVNVERLGRGIETGTANSILVKVNOIGTLTETTFDAVDMHRAAG 360
QY 356 YTSMSHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYA 415
Db 361 YTSMSHRSGETEDTTIADLAVALNCGQIKTGASASRSDRVAKYNOLLRIEGLGAAARYA 420
QY 416 GRSF 420
Db 421 GKSF 425

RESULT 4
US-09-710-279-2222
; Sequence 2222, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2222
; LENGTH: 434
; TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2222

Query Match      60.6%; Score 1306.5; DB 4; Length 434;
Best Local Similarity 60.8%; Pred. No. 4.4e-112;
Matches 259; Conservative 64; Mismatches 94; Indels 9; Gaps 3;

QY 4 IMHVAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 62
Db 4 ITDVYAREVLDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 63

QY 63 VLKAVENVNEEIGDELAGLEAD--DORLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 120
Db 64 VTKAVENVNEEIAPEIVEGEFSVLDQVSDIKMQLDGTGTHNKGKLGANAILGVSIARA 123

QY 121 AADSAGLPLFRYIGGPNNAHVLPVPMNIIINGGAHADSGVDVQEFMTAPIGAETFSBALRN 180
Db 124 AADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQBFMILPVGAESFKESLRW 183

QY 181 GAENVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Db 184 GAEIFHNLSILSERGLETAVGDEGGFAPRFEGETDAVETIIKAIKAGYKPGEDVFLGF 243

QY 241 DVASSEFFKDGTYHF-----EGQHSAAEMANVYAEVLVDAYPIVSTEDPLQDDWEGYT 294
Db 244 DCASSEFFYENGVDYTKFGEHGAKSAAEQVDVLEELIGKYPIIITIEDGMDENDWEGWK 303

QY 295 NLTATIGDKVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTTFDADVMAHRA 354
Db 304 QLTDRIGDKVQLVGDDLFVTNTEILSKIEQIGNSILIKVNOIGTLTTFDADAEMAQA 363

QY 355 GYTSMMSHRSGETETTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVY 414
Db 364 GYTAVVSHRSGETETTTIADIAVATNAGQIKTGSLSRTDRIAKYNOLLRIEDELSETAKF 423

QY 415 AGRSAF 420
Db 424 EGIKSP 429

RESULT 5
US-09-134-001C-5641
; Sequence 5641, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT FILING DATE: 1997-08-14
; PRIOR FILING DATE: 1997-08-14
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5641
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5641

Query Match      60.6%; Score 1306.5; DB 3; Length 440;
Best Local Similarity 60.8%; Pred. No. 4.5e-112;
Matches 259; Conservative 64; Mismatches 94; Indels 9; Gaps 3;

QY 4 IMHVAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 62
Db 10 ITDVYAREVLDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 69

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-2222

Query Match      60.6%; Score 1306.5; DB 4; Length 434;
Best Local Similarity 60.8%; Pred. No. 4.4e-112;
Matches 259; Conservative 64; Mismatches 94; Indels 9; Gaps 3;

QY 4 IMHVAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 62
Db 4 ITDVYAREVLDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 63

QY 63 VLKAVENVNEEIGDELAGLEAD--DORLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 120
Db 64 VTKAVENVNEEIAPEIVEGEFSVLDQVSDIKMQLDGTGTHNKGKLGANAILGVSIARA 123

QY 121 AADSAGLPLFRYIGGPNNAHVLPVPMNIIINGGAHADSGVDVQEFMTAPIGAETFSBALRN 180
Db 124 AADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQBFMILPVGAESFKESLRW 183

QY 181 GAENVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Db 184 GAEIFHNLSILSERGLETAVGDEGGFAPRFEGETDAVETIIKAIKAGYKPGEDVFLGF 243

QY 241 DVASSEFFKDGTYHF-----EGQHSAAEMANVYAEVLVDAYPIVSTEDPLQDDWEGYT 294
Db 244 DCASSEFFYENGVDYTKFGEHGAKSAAEQVDVLEELIGKYPIIITIEDGMDENDWEGWK 303

QY 295 NLTATIGDKVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTTFDADVMAHRA 354
Db 304 QLTDRIGDKVQLVGDDLFVTNTEILSKIEQIGNSILIKVNOIGTLTTFDADAEMAQA 363

QY 355 GYTSMMSHRSGETETTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLGDAGVY 414
Db 364 GYTAVVSHRSGETETTTIADIAVATNAGQIKTGSLSRTDRIAKYNOLLRIEDELSETAKF 423

QY 415 AGRSAF 420
Db 424 EGIKSP 429

RESULT 6
US-09-830-217-22
; Sequence 22, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461PCT
; CURRENT APPLICATION NUMBER: US/09/830,217
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-830-217-22

Query Match      59.4%; Score 1280.5; DB 4; Length 434;
Best Local Similarity 59.1%; Pred. No. 1.1e-109;
Matches 253; Conservative 68; Mismatches 94; Indels 13; Gaps 4;

QY 4 IMHVAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 62
Db 4 ITDVYAREVLDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELDRGG--DRYLKGG 63

QY 63 VLKAVENVNEEIGDELAGLEAD--DORLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 118
Db 64 VTKAVENVNEEIAPEI--IEGFSVLDQVSDIKMQLDGTGTHNKGKLGANAILGVSIARA 121

QY 119 KAAASAGLPLFRYIGGPNNAHVLPVPMNIIINGGAHADSGVDVQEFMTAPIGAETFSBAL 178
Db 122 RAAADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQBFMILPVGATTFKESL 181

QY 179 RGAENVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIAL 238
Db 182 RWGTEIFHNLSILSQRLGLETAVGDEGGFAPKPEGETDAVETIIQIAEAGYKPGBEVFL 241

```



```

; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3810:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...398
; SEQUENCE DESCRIPTION: SEQ ID NO: 3810:
US-09-107-532A-3810

Query Match 57.5%; Score 1240; DB 4; Length 398;
Best Local Similarity 62.1%; Pred. No. 5.3e-106;
Matches 242; Conservative 59; Mismatches 83; Indels 6; Gaps 3;

QY 37 VPFGASTGVHEAHELDDGGD-RYLGGKVLKAVNNVNEEIGDELAGLEADDORLIDEAMIK 95
DB 3 VPFGASTGEYAEVELDDGDKARYGGKGVTKAVDNNVNIATAEIIIGYDVRDQMAIDKAMIA 62
QY 96 LDGTANKSRLGANAIIIGVSMVAKAAADSAGLPFLFRYIGGPNNAHVLPVPMNIIINGGAHA 155
DB 63 LDGTPNKGKLGANAIIIGVSIVARAAADYLEVLYHYLGGFNTKVLPTPMNIIINGGSHA 122
QY 156 DSGVDVQEFMIAPIGAIETSEALRNGAEVTHALKSVIKKGLSTGLGDECGFAPSVCSTR 215
DB 123 DNSIDPQEFMIPVGAPTTFKEALRMGAEPFHALLASILKARGLATSVGDEGFPAPNLGSNE 182
QY 216 EALDLTVKAIKAGFTPGKDIALALDVASSEFF--KDGTVHFEG--GOHSAEAEMANVYA 270
DB 183 EGFEVITAEIRKAGYVPGKDVLLMDAAASEFTDKEKGVTVLADSGEGEKTITDEMIKFEY 242
QY 271 ELVDAYPIVSIEDPLEQDMEGYTNLTATIGDKVQIVGDDFFVTNPNRLKEGIAKKAANS 330
DB 243 ELVSKYPIISIEDGLDENDWDGFKLTDVLGDKVQLVGDDLFVTNTQKLSEGIEKGIAN 302
QY 331 ILVKVNIQIGTLTETPDVDMAHRAGYTSMMSHRSGETEDTTIADLAVLNCGOIKTGAPA 390
DB 303 ILIKVNIQIGTLTETFEAIEWAKEAGYATVVSHRSGETEDSTISDIAVATNAGQIKTGSL 362
QY 391 RSDRAKYNOLLRIEQLLDGAGVYAGRSF 420
DB 363 RTDRIAKYNOLLRIEQLDGEVAEYKGLKGSF 392

RESULT 11
US-09-328-352-7269
; Sequence 7269, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7269
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7269

```

```

Query Match          57.1%; Score 1231.5; DB 4; Length 444;
Best Local Similarity 59.3%; Pred. No. 3.8e-105;
Matches 253; Conservative 65; Mismatches 100; Indels 9; Gaps 5;

QY 1 VAEIMHVAFRILDSRGNPTVEAEVFLDDSGHGVAGVPSGASTGVHEAHLRDGG-DRYL 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
16 MSQIVDIRAREILDSRGNPTIEADVILESGVVGACAPSGASTGREALRLDGDGKSYL 75
QY 60 GKGVLKAVENVEEIGDELGLAGLEADDQRLIDEAMIKLDTGANKSRILGANAILGVSMVAK 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
76 GKGVRTAVQNVNNSIHLLVQSVFEQKALDERKMFADGTENKSKLGNATLAVSLAAAH 135
QY 120 AAADSAGLPLFRYITG---GPNHVLVPMNININGGAHADSGVDVQSFMIAPIGAEFTSE 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 AAAAEQKLPFLQYIANLRGQTTLMPVPMNINLNGGAHADNTVDIQEFMIEPVGFTSFAE 195
QY 177 ALRNGAEVYHALKSVIEKGLSTGLDGGFAPSVGSTRALDLIVKAEKAGFTPGKOI 236
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 ALRAGAEVHSLKSVLKQGLNTAVGDEGGFAPNLRSEBAITVILQAEQTYKAGSDI 255
QY 237 ALALDVASSBFFXGTYHFEQ-GQHS--AAEMANVVAELVDVPIVSIEDPLEDDWEGY 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
256 MLALDCASSBFFXNGQYILSEGNKSFSTNQPADYLAGLVKQPIIISIEDGDESDEWG 315
QY 294 TNLTAITGDKVQVGGDFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHR 353
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
316 SYLTSILGDKIQLVGGDLFVTNPKILQRIGIDEKVGNSILIKYNIQIGTLTETDLAIVLAKA 375
QY 354 AGYTSMMSHRSGETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNQLLRIEQLLDGAV 413
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
376 NGYTVIHSRSGETEDTTIADLAVGTAAGQIKTGSICRSRVSKYNQLLRIEEL--TKAV 433
QY 414 YAGRSF 420
Db 434 YRGKAEF 440

```

```

RESULT 12
US-08-847-065-19
; Sequence 19, Application US/08847065
; Patent No. 6245335
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wizemann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,065
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Seq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-158 ..
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: B. subtilis
; US-08-847-065-19

```

```

Query Match          56.9%; Score 1226.5; DB 3; Length 420;
Best Local Similarity 58.7%; Pred. No. 1e-104;
Matches 244; Conservative 61; Mismatches 100; Indels 11; Gaps 4;

QY 15 SRGNPTVEAEVFLDDSGHGVAGVPSGASTGVHEAHLRDGG-DRYLKGLKAVENVEE 73
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 SRGNPTLEVEVTESAFGRGMVPSGASTGEHEAHLRDGDKSRYGGLGTQKAVDNVNNI 60
QY 74 IGDELAGLEADDQRLIDEAMIKLDTGANKSRILGANAILGVSMVAKAAADASAGLPLFRYI 133
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 IADSIIGYDVRDQADRAMIALDGTGPNKGLGANAILGVSIAPFAAADYLEIPLYSYL 120
QY 134 GGNHVLVPMNININGGAHADSGVDVQSFMIAPIGAEFTSEALRNGAEVYHALKSVIK 193
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 GGFNTKVLPTPMNININGGSHSDAPAFQEFMILPVGAPTFKETLRYGAEIHALKKILK 180
QY 194 EKSLSTGLDGGFAPSVGSTRALDLIVKAEKAGFTPGKDIALDALDVASSEFFK--- 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 SRGLETAVDEGGFAPRPEGTEDGVETILAAIEAGYVPGKDVFLGFCASSEFYDKERK 240
QY 250 ---DGYHFEQ---GQHSAAEMANVVAELVDVPIVSIEDPLEDDWEGYTNLTATIGDKV 304
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 VYDYT-KFEGEAAVRTSAEQIDYLEELVKNYPIITIEDGMDENDWDGKALTERLGKKV 299
QY 305 QIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGTLTETFDVADMAHRPAGYTSMMSHRS 364
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 QLVGDDFFVTNTDYLARGIQKGAANSILIKVNIQIGTLTETFEAIEAKEAGYTAVVSHRS 359
QY 365 GETEDTTIADLAVALNCGOIKTGAPARSDRVAKYNQLLRIEQLLDGAVYAGRSF 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
360 GETEDSTIADIATNAGQIKTGSLSRTDRIAKYNQLLRIEQLGEVAEYRGLKSF 415

```

```

RESULT 13
US-09-829-382-19
; Sequence 19, Application US/09829382
; Patent No. 6784164
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wizemann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 12:13:58 ; Search time 1026 Seconds
(without alignments)
171.954 Million cell updates/sec

Title: US-10-728-947-4

Perfect score: 2155

Sequence: 1 VAEIMHVFAREILDSRGNT.....QLLGDAVYAGRSAPFRFGQ 425

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11_NEW PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2155	100.0	425	9	US-09-860-768-4
2	2155	100.0	425	16	US-10-728-947-4
3	2151	99.8	425	9	US-09-860-768-2
4	2151	99.8	425	16	US-10-728-947-2
5	2148	99.5	425	9	US-09-738-626-4585
6	2145	99.5	425	16	US-10-781-014-72
7	2145	99.5	425	17	US-10-494-836-8
8	1913	88.8	425	15	US-10-282-122A-53908
9	1562	72.5	423	15	US-10-369-493-8340
10	1560.5	72.4	429	15	US-10-282-122A-62133
11	1553.5	72.1	429	15	US-10-282-122A-64490

12	1553.5	72.1	439	15	US-10-282-122A-62696	Sequence 62696, A
13	1551.5	72.0	428	14	US-10-156-761-11070	Sequence 11070, A
14	1502.5	69.7	447	15	US-10-282-122A-63682	Sequence 63682, A
15	1422	66.0	423	15	US-10-369-493-11696	Sequence 11696, A
16	1412.5	65.5	426	15	US-10-369-493-16911	Sequence 16911, A
17	1402.5	65.1	423	15	US-10-369-493-10557	Sequence 10557, A
18	1402.5	65.1	426	15	US-10-369-493-17886	Sequence 17886, A
19	1402.5	65.1	429	15	US-10-369-493-17433	Sequence 17433, A
20	1396	64.8	424	15	US-10-369-493-11819	Sequence 11819, A
21	1394	64.7	420	15	US-10-369-493-14739	Sequence 14739, A
22	1394	64.7	423	15	US-10-369-493-15222	Sequence 15222, A
23	1392.5	64.6	431	15	US-10-282-122A-52760	Sequence 52760, A
24	1377.5	63.9	430	15	US-10-369-493-23299	Sequence 23299, A
25	1373	63.4	431	15	US-10-282-122A-46167	Sequence 46167, A
26	1366.5	63.4	448	15	US-10-282-122A-52925	Sequence 52925, A
27	1362.5	63.2	426	15	US-10-369-493-21730	Sequence 21730, A
28	1362.5	63.2	427	15	US-10-369-493-21006	Sequence 21006, A
29	1359.5	63.1	430	15	US-10-282-122A-60954	Sequence 60954, A
30	1357.5	63.0	430	15	US-10-369-493-13491	Sequence 13491, A
31	1354.5	62.9	431	15	US-10-282-122A-51920	Sequence 51920, A
32	1351.5	62.7	427	15	US-10-282-122A-47804	Sequence 47804, A
33	1347.5	62.5	417	15	US-10-369-493-9097	Sequence 9097, Ap
34	1347.5	62.5	424	15	US-10-369-493-7762	Sequence 7762, Ap
35	1347.5	62.5	426	15	US-10-369-493-4688	Sequence 4688, Ap
36	1347.5	62.5	427	15	US-10-369-493-7447	Sequence 7447, Ap
37	1347.5	62.5	427	15	US-10-282-122A-49282	Sequence 49282, A
38	1342.5	62.3	428	16	US-10-375-266-38	Sequence 38, Appl
39	1342	62.3	429	9	US-09-815-242-5012	Sequence 5012, Ap
40	1342	62.3	432	9	US-09-815-242-10854	Sequence 10854, A
41	1342	62.3	432	15	US-10-282-122A-42554	Sequence 42554, A
42	1334.5	61.9	426	15	US-10-369-493-32	Sequence 32, Appl
43	1332.5	61.8	428	15	US-10-369-493-9970	Sequence 9970, Ap
44	1327.5	61.6	424	15	US-10-369-493-20662	Sequence 20662, A
45	1322	61.3	422	15	US-10-369-493-10435	Sequence 10435, A

ALIGNMENTS

RESULT 1
US-09-860-768-4
; Sequence 4, Application US/09860768
; Patent No. US20020082403A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Bathe, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/09/860,768
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-860-768-4

Query Match 100.0%; Score 2155; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 6.7e-171;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHLRDGDRYLG	60
DB	1	VAEIMHVFAREILDSRGNTVFAEVLDDGSHGVAGVPSGASTGVHEAHLRDGDRYLG	60
QY	61	KGVLKAVENVNNEIGDELGLAGLADDDORLIIDEAMI KLDGTANKSRRLGANAILGVSMVAKA	120
DB	61	KGVLKAVENVNNEIGDELGLAGLADDDORLIIDEAMI KLDGTANKSRRLGANAILGVSMVAKA	120

QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 |||||
 Db 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 |||||
 QY 181 GAEVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 |||||
 Db 181 GAEVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 |||||
 QY 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||
 Db 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||
 QY 301 GDKVQIVGDDPFTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 360
 |||||
 Db 301 GDKVQIVGDDPFTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 360
 |||||
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRLEQLLDGAGVYAGRS 420
 |||||
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRLEQLLDGAGVYAGRS 420
 |||||
 QY 421 PRFQ 425
 |||||
 Db 421 PRFQ 425
 |||||

RESULT 2

US-10-728-947-4
 ; Sequence 4, Application US/10728947
 ; Publication No. US2004020394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/10/728,947
 ; PRIOR FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US/09/860,768
 ; PRIOR FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-10-728-947-4

Query Match 100.0%; Score 2155; DB 16; Length 425;
 Best Local Similarity 100.0%; Pred. No. 6.7e-171;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 |||||
 Db 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 |||||
 QY 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 120
 |||||
 Db 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 120
 |||||
 QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 |||||
 Db 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 |||||
 QY 181 GAEVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 |||||
 Db 181 GAEVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 |||||
 QY 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||
 Db 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||

Db 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||
 QY 301 GDKVQIVGDDPFTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 360
 |||||
 Db 301 GDKVQIVGDDPFTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 360
 |||||
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRLEQLLDGAGVYAGRS 420
 |||||
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRLEQLLDGAGVYAGRS 420
 |||||
 QY 421 PRFQ 425
 |||||
 Db 421 PRFQ 425
 |||||

RESULT 3
 US-09-860-768-2
 ; Sequence 2, Application US/09860768
 ; Patent No. US20020082403A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; APPLICANT: Hermann, Thomas
 ; APPLICANT: Pohler, Alfred
 ; APPLICANT: Kalinowski, Jorn
 ; APPLICANT: Bathe, Brigitte
 ; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
 ; FILE REFERENCE: 21123/278404
 ; CURRENT APPLICATION NUMBER: US/09/860,768
 ; CURRENT FILING DATE: 2001-05-21
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 425
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-860-768-2

Query Match 99.8%; Score 2151; DB 9; Length 425;
 Best Local Similarity 99.8%; Pred. No. 1.5e-170;
 Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 |||||
 Db 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHELKGGDRYLG 60
 |||||
 QY 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 120
 |||||
 Db 61 KGVLKAVENNVNEEIGDELAGLEADDQRLIDEAMI KLDGTANKSRLGANAILGVSMVAKA 120
 |||||
 QY 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 |||||
 Db 121 AADSAGLPLFRYIGGPNNAHVLVPMNMNIINGGAHADSGVDVQEFMIAPIGAETPSEALRN 180
 |||||
 QY 181 GAEVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 |||||
 Db 181 GAEVYHALKSVIKEKGLSTGLDGGGFPAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
 |||||
 QY 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||
 Db 241 DVASSEFFKQGYHFEFGGQHSAAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATI 300
 |||||
 QY 301 GDKVQIVGDDPFTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 360
 |||||
 Db 301 GDKVQIVGDDPFTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM 360
 |||||
 QY 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRLEQLLDGAGVYAGRS 420
 |||||
 Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRLEQLLDGAGVYAGRS 420
 |||||
 QY 421 PRFQ 425
 |||||
 Db 421 PRFQ 425
 |||||

RESULT 4
US-10-728-947-2
; Sequence 2, Application US/10728947
; Publication No. US2004020394A1
; GENERAL INFORMATION:
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; APPLICANT: Hermann, Thomas
; APPLICANT: Pohler, Alfred
; APPLICANT: Kalinowski, Jörn
; APPLICANT: Bache, Brigitte
; TITLE OF INVENTION: New Nucleotide Sequences that Code for the Eno Gene
; FILE REFERENCE: 21123/278404
; CURRENT APPLICATION NUMBER: US/10728,947
; PRIOR FILING DATE: 2003-12-08
; PRIOR APPLICATION NUMBER: US/09/860,768
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-728-947-2

Query Match 99.8%; Score 2151; DB 16; Length 425;
Best Local Similarity 99.8%; Pred. No. 1.5e-170; Indels 0; Gaps 0;
Matches 424; Conservative 1; Mismatches 0;

Qy	1	VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG	60
Db	1	VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG	60
Qy	61	KGVLKAVENNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA	120
Db	61	KGVLKAVENNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA	120
Qy	121	AADSAGLPFRYIGGPNHVLVPMNIIINGGAHDSGVQVDFPMIPIGAETTFSEALRN	180
Db	121	AADSAGLPFRYIGGPNHVLVPMNIIINGGAHDSGVQVDFPMIPIGAETTFSEALRN	180
Qy	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL	240
Db	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL	240
Qy	241	DVASSEPFKDGTYHFEGGQSAEMANVYAEVDAIPVISEDPLQEDDWEYTNLTATI	300
Db	241	DVASSEPFKDGTYHFEGGQSAEMANVYAEVDAIPVISEDPLQEDDWEYTNLTATI	300
Qy	301	GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM	360
Db	301	GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM	360
Qy	361	SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF	420
Db	361	SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF	420
Qy	421	PRFQG 425	
Db	421	PRFQG 425	

RESULT 5
US-09-738-626-4585
; Sequence 4585, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4585
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4585

Query Match 99.7%; Score 2148; DB 9; Length 425;
Best Local Similarity 99.5%; Pred. No. 2.6e-170; Indels 0; Gaps 0;
Matches 423; Conservative 2; Mismatches 0;

Qy	1	VAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG	60
Db	1	MAEIMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHLRGGDRYLG	60
Qy	61	KGVLKAVENNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA	120
Db	61	KGVLKAVENNEEIGDELADQRLIDEAMIKLDGTANKSRGLGANAILGVSMVAKA	120
Qy	121	AADSAGLPFRYIGGPNHVLVPMNIIINGGAHDSGVQVDFPMIPIGAETTFSEALRN	180
Db	121	AADSAGLPFRYIGGPNHVLVPMNIIINGGAHDSGVQVDFPMIPIGAETTFSEALRN	180
Qy	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL	240
Db	181	GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAEKAGTTPGKDIALAL	240
Qy	241	DVASSEPFKDGTYHFEGGQSAEMANVYAEVDAIPVISEDPLQEDDWEYTNLTATI	300
Db	241	DVASSEPFKDGTYHFEGGQSAEMANVYAEVDAIPVISEDPLQEDDWEYTNLTATI	300
Qy	301	GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM	360
Db	301	GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNOIGTLTETFDVADMAHRAGYTSM	360
Qy	361	SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF	420
Db	361	SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGVYAGRSF	420
Qy	421	PRFQG 425	
Db	421	PRFQG 425	

RESULT 6
US-10-781-014-72
; Sequence 72, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kröger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zeider, Oskar
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; INVOLVED IN CARBON METABOLISM AND ENERGY

```

; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: BGI-126CFN
; CURRENT APPLICATION NUMBER: US/10/781.014
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/151,572
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19931412.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931424.1
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931428.4
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 784
; SEQ ID NO 72
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-781-014-72

```

```

Query Match          99.5%; Score 2145; DB 16; Length 425;
Best Local Similarity 99.5%; Pred. No. 4.6e-170;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Db 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Qy 61 KGVLKAVENNVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120
Qy 121 AADSAGLPLFRYTGGNNAHVLVPMNNIINGGAHDSGVVDVQBFMTAPIGAETTFSEALRN 180
Db 121 AADSAGLPLFRYTGGNNAHVLVPMNNIITGGAHDSGVVDVQBFMTAPIGAETTFSEALRN 180
Qy 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Db 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Qy 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAPIVSIEDPLQEDDWEGETNLAT 300
Db 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAPIVSIEDPLQEDDWEGETNLAT 300
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHAGYTSMM 360
Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHAGYTSMM 360
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGYAGRSF 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGYAGRSF 420
Qy 421 PRFQG 425
Db 421 PRFQG 425

```

```

RESULT 7
US-10-494-836-8
; Sequence 8, Application US/10494836
; Publication No. US20050014233A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Hartwig
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopffrogge, Corinna
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy products
; FILE REFERENCE: BGI-1670US
; CURRENT APPLICATION NUMBER: US/10/494,836
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: PCT/EP02/12135
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 101 54 270.4
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 8
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-494-836-8

```

```

Query Match          99.5%; Score 2145; DB 17; Length 425;
Best Local Similarity 99.5%; Pred. No. 4.6e-170;
Matches 423; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Db 1 VAEIMHVFAREILDSRGNPVTEAEVFLDDGSHGVAGVPSGASTGVHEAHLRDGGDRYLG 60
Qy 61 KGVLKAVENNVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120
Db 61 KGVLKAVENNVNEIGDELADDDQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKA 120
Qy 121 AADSAGLPLFRYTGGNNAHVLVPMNNIINGGAHDSGVVDVQBFMTAPIGAETTFSEALRN 180
Db 121 AADSAGLPLFRYTGGNNAHVLVPMNNIITGGAHDSGVVDVQBFMTAPIGAETTFSEALRN 180
Qy 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Db 181 GAEVYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLIVKAIKAGFTPGKDIALAL 240
Qy 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAPIVSIEDPLQEDDWEGETNLAT 300
Db 241 DVASSEFFKDGTVHFEGGQSHAAEMANVYAEVDAPIVSIEDPLQEDDWEGETNLAT 300
Qy 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHAGYTSMM 360
Db 301 GDKVQIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQIGLTETTFDAVDMHAGYTSMM 360
Qy 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGYAGRSF 420
Db 361 SHRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNOLLRIEQLLDGAGYAGRSF 420
Qy 421 PRFQG 425
Db 421 PRFQG 425

```

```

RESULT 8
US-10-282-122A-53908
; Sequence 53908, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangou
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

```



```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 62133
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-62133

```

```

Query Match 72.4%; Score 1560.5; DB 15; Length 429;
Best Local Similarity 72.9%; Pred. No. 2.6e-121;
Matches 307; Conservative 47; Mismatches 66; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
Db 4 IEQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
QY 64 LKAVENNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGTPDKSRGNGAILGVSLAVAKAAAD 123
QY 124 SAGLPLFRYIGGPNNAHLVPVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSBALRWGAE 183
Db 124 SAGLPLFRYIGGPNNAHLVPVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSBALRWGAE 183
QY 184 VYHALKSVIKEGLSTGLDGGGFPAPSVGSTRALDLIVKAIKAGFTPGKDIALLDVA 243
Db 184 VYHLSKSVLKKEGLSTGLDGGGFPAPSVGSTRALDLIVKAIKAGFTPGKDIALLDVA 243
QY 244 SSEFFPKDGT-YHFEQGGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATIGD 302
Db 244 ATEFYSDGTGYKEGSTRTEAEQNAEFYAGLLGAYPIVSIEDPLQEDDWEQYTNLTATIGD 303
QY 303 KVOIVGDDFFVTNPERLKEGIAKKAANSILVKVNIQTGLTETPDVDMRAHAGYTSMSH 362
Db 304 RVQLVGDDVFTNPERLEEGIEKGVANALLVKVNIQTGLTETLDAVALAHSSGYRTMSH 363
QY 363 RSETEDTTADIADVNCQIKTGAPASDRVAKVKNQILRTEQLLGDAGVAGVAGSAPFR 422
Db 364 RSETEDTTADIADVAVGSGQIKTGAPASDRVAKVKNQILRTEQLLGDAGVAGVAGSAPFR 423
QY 423 F 423
Db 424 F 424

```

```

RESULT 11
US-10-282-122A-64490
; Sequence 64490, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64490
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64490

```

```

Query Match 72.1%; Score 1553.5; DB 15; Length 429;
Best Local Similarity 73.2%; Pred. No. 9.8e-121;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

QY 4 IMHVFAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
Db 4 IEQVGAREILDSRGNTPTVEAEVFLDDGSHGVAGVPSGASTGVHEAHELRDGGDRYLKGV 63
QY 64 LKAVENNEEIGDELGLAGLEADDQRLIDEAMIKLDGTANKSRIGANAILGVSMVAKAAAD 123
Db 64 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLDTGTPDKSRGNGAILGVSLAVAKAAAD 123
QY 124 SAGLPLFRYIGGPNNAHLVPVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSBALRWGAE 183
Db 124 SAGLPLFRYIGGPNNAHLVPVPMNIIINGGAHDSGVVDVQEFMTAPIGAETFSBALRWGAE 183
QY 184 VYHALKSVIKEGLSTGLDGGGFPAPSVGSTRALDLIVKAIKAGFTPGKDIALLDVA 243
Db 184 VYHLSKSVLKKEGLSTGLDGGGFPAPSVGSTRALDLIVKAIKAGFTPGKDIALLDVA 243
QY 244 SSEFFPKDGT-YHFEQGGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEQYTNLTATIGD 302

```

```
Db 244 ATEFFDTGTVFEGTTRTADQMTFVAGLIGAYPLVSIEDPLSEDDWGDGWAALTASIGD 303
Qy 303 KVOIVGDDFFVTNPERLKEGIKAANSILVKVNOIGTLTETPDVDMHRAGYTSMWSH 362
Db 304 RVQIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETPDVLAHGGYRTMISH 363
Qy 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 422
Db 364 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEALGDAAARYAGDLAPPR 423
Qy 423 F 423
Db 424 F 424
```

RESULT 12

```
US-10-282-122A-62696
; Sequence 62696, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-06-28
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62696
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62696
```

```
Query Match 72.1%; Score 1553.5; DB 15; Length 439;
Best Local Similarity 73.2%; Pred. No. 1e-120;
Matches 308; Conservative 42; Mismatches 70; Indels 1; Gaps 1;

Qy 4 IMHVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGGDRYLKGV 63
Db 14 IEQVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGGDRYLKGV 73
```

```
Qy 64 LKAVENVNEBIGDELADQRLIDEAMIKLDGTANKSRLGANAILGVSMVAKAAAD 123
Db 74 QKAVQAVLDEIGPAVIGLNADDQRLVDQALVDLGDGTPDKSRLGGNAILGVSLAVAKAAAD 133
Qy 124 SAGLPLFRYGGNAHVLPVPMNNIINGGAHAGSVGVQEFMTAPIGAETTFSEALRNGAE 183
Db 134 SABLPLFRYVGGNAHVLPVPMNNIINGGAHADTAVDIOEFMTAPIGAPSFVEALRNGAE 193
Qy 184 VYHALKSVIKEKGLSTGLGDEGGFAPSVGSTREALDLI VKAIEKAGFTPGKDIALALDVA 243
Db 194 VYHALKSVLKEKGLSTGLGDEGGFAPDVAGTTAALDLISRAIESAGURPGADVALALDAA 253
Qy 244 SSEFFKDGTT-YHFEQGGHSAEMANVYAEVLVDAYPIVSIEDPLQEDDWEGYTNLTATIGD 302
Db 254 ATEFFDTGTVFEGTTRTADQMTFVAGLIGAYPLVSIEDPLSEDDWGDGWAALTASIGD 313
Qy 303 KVOIVGDDFFVTNPERLKEGIKAANSILVKVNOIGTLTETPDVDMHRAGYTSMWSH 362
Db 314 RVQIVGDDIFVTNPERLEEGIERGVANALLVKVNOIGTLTETPDVLAHGGYRTMISH 373
Qy 363 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLLGDAGVYAGRSAPPR 422
Db 374 RSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEALGDAAARYAGDLAPPR 433
Qy 423 F 423
Db 434 F 434
```

RESULT 13

```
US-10-156-761-11070
; Sequence 11070, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11070
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11070
```

```
Query Match 72.0%; Score 1551.5; DB 14; Length 428;
Best Local Similarity 72.9%; Pred. No. 1.4e-120;
Matches 310; Conservative 43; Mismatches 71; Indels 1; Gaps 1;

Qy 1 VAEIMEVFAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGGDRYL 59
Db 1 VPSIDVVVAREILDSRGNPTVEAEVFLDDGSHGVAGVPSGASTGVHAEHLRGGDRYL 60
Qy 60 KGKVLKAVENVNEBIGDELADQRLIDEAMIKLDGTANKSRLGANAILGVSMVAK 119
Db 61 KGKVLKAVLAVIEQIGPELVGYDATEQRLIDQAMFDLDTDNKGLGANAILGVSLAVAH 120
Qy 120 AAADSAGLPLFRYGGNAHVLPVPMNNIINGGAHADSGVDVQEFMTAPIGAETTFSEALR 179
Db 121 AAASEADPLFRYIGGPNHLLPVPMMNILINGGSHADSNVDIOEFMTAPIGAESFSEALR 180
```

QY 180 NGAEVYHALKSVIKGLSTGLDGGGFPVSGVSTREALDLIVKAIKAGFTPGKDIALA 239
 Db 181 WGAEVYHTLKKVLTGKSTGLDGGGFPVSGVSTREALDLIVKAIKAGFTPGKDIALA 240
 QY 240 LDVASSEFFKDGTYHFGGQHSAAEMANVYAEVLDVAYPIVSIEDPLQEDDWEQYTNLTAT 299
 Db 241 LDVAASEFYKDGKVEFGKSRSAEMTEYYEELVSAVPLVSIEDPLVEDDWDAGWKVITDK 300
 QY 300 IGDVQVIVGDDFFVTNPERLKEGIKAAKANSILVKNQIQTTLTETFDVDMHRAAGVTSM 359
 Db 301 LGDKVQIVGDDFFVTNPERLARGIEGSAEALVKNQIQTTLTETFDVDMHRAAGVTSM 360
 QY 360 MSHRSGTETDTTADLAVANCGQIKTGAPARSDRVAKYNQIQTTLTETFDVDMHRAAGVTSM 419
 Db 361 MSHRSGTETDTTADLAVANCGQIKTGAPARSDRVAKYNQIQTTLTETFDVDMHRAAGVTSM 420
 QY 420 FPRFQ 424
 Db 421 FPRF 425

RESULT 14
 US-10-282-122A-63682
 ; Sequence 63682, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zykand, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: EUTRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 63682
 ; LENGTH: 447
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium leprae
 US-10-282-122A-63682

Query Match 69.7%; Score 1502.5; DB 15; Length 447;
 Best Local Similarity 70.8%; Pred. No. 1.9e-116;

Matches 300; Conservative 47; Mismatches 76; Indels 1; Gaps 1;
 QY 1 VAEIMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVVHEAHELDDGDRYLG 60
 Db 19 VPVIEQVGAAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVVHEAHELDDGDRYLG 78
 QY 61 KGVLKAVENVNEEIGDELADDDORLIDEAMIKLDGTANKSRGALGANAILGVSMVAKA 120
 Db 79 KGVKRAVDAVLDIGPWWIQLANDQRLIDQELLDDGTDPKSRGALGANAILGVSLAVAKA 138
 QY 121 AADSAGLPLFRYIGGNNAHVLPVPMNIIINGGAHDSGVVDQEFMTAPIGAETFSBALRN 180
 Db 139 AADSAGLPLFRYIGGNNAHVLPVPMNIIINGGAHDSGVVDQEFMTAPIGAETFSBALRN 198
 QY 181 GAEVYHALKSVIKGLSTGLDGGGFPVSGVSTREALDLIVKAIKAGFTPGKDIALA 240
 Db 199 GAEVYHALKSVIKGLSTGLDGGGFPVSGVSTREALDLIVKAIKAGFTPGKDIALA 258
 QY 241 DVASSEFFKDGTYHFGGQHSAAEMANVYAEVLDVAYPIVSIEDPLQEDDWEQYTNLTAT 299
 Db 259 DAAATEFYTDGIGYHFGGQHSAAEMANVYAEVLDVAYPIVSIEDPLQEDDWEQYTNLTAT 318
 QY 300 IGDVQVIVGDDFFVTNPERLKEGIKAAKANSILVKNQIQTTLTETFDVDMHRAAGVTSM 359
 Db 319 IGEQVQIVGDDFFVTNPERLKEGIKAAKANSILVKNQIQTTLTETFDVDMHRAAGVTSM 378
 QY 360 MSHRSGTETDTTADLAVANCGQIKTGAPARSDRVAKYNQIQTTLTETFDVDMHRAAGVTSM 419
 Db 379 ISHRSGTETDTTADLAVANCGQIKTGAPARSDRVAKYNQIQTTLTETFDVDMHRAAGVTSM 438
 QY 420 FPRF 423
 Db 439 FURY 442

RESULT 15
 US-10-369-493-11696
 ; Sequence 11696, Application US/10369493
 ; Publication No. US2003033675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 11696
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Agrobacterium tumefaciens
 US-10-369-493-11696

Query Match 66.0%; Score 1422; DB 15; Length 423;
 Best Local Similarity 66.2%; Pred. No. 8.9e-110;
 Matches 276; Conservative 58; Mismatches 81; Indels 2; Gaps 1;
 QY 4 IMHVFAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVVHEAHELDDGDRYLGKGV 63
 Db 4 ITDIIAREILDSRGNTVEAEVFLDDGSHGVAGVPSGASTGVVHEAHELDDGDRYLGKGV 63
 QY 64 LKAVENVNEEIGDELADDDORLIDEAMIKLDGTANKSRGALGANAILGVSMVAKA 123
 Db 64 EKAVEAVNTEIFDAIGGFAENQIQIDQMIALDGTANKSRGALGANAILGVSMVAKA 123
 QY 124 SAGLPLFRYIGGNNAHVLPVPMNIIINGGAHDSGVVDQEFMTAPIGAETFSBALRN 183

Db	124	ASGLPLYRYVGGPNAHLLPVPMMNIINGGAHADNPIDQEFMILPVGAEINREAVRMGE	183
Qy	184	VYHALKSVIKEKGLSTGLGDEGGFAPSVGSTRALDLIVKAIKAGETPCGDIALALDVA	243
Db	184	VFTLKKELSAQHNTNVGDEGGFAPGLESAPALDFIMKSIKAGYRPGEDMYVGLDCA	243
Qy	244	SSEFFKDGTYHFEGQHS--AAEMANYAELVDAYPIVSIEDPLQEDDWEGYTNLTATIG	301
Db	244	STEFFKDGKYLEGEGRTEPGAMAEYLAELVNKYPIISVEDGMAEDDWEGWKTLDLVG	303
Qy	302	DKVQIVCDDDFVTNPERLKEGLAKKAANSILVKVNOIGTLTETFDVDMAHHRAGYTSMS	361
Db	304	NKCQLVGDDLFVTNSAKLRDGIKXGVANSILVKVNOIGSLSETLDAVETAHKAGYTAVMS	363
Qy	362	HRSGETEDTTIADLAVALNCGQIKTGAPARSDRVAKYNQLLRIEQLIGDAGVYAGRS	418
Db	364	HRSGETEDSTIADLAVATNCGQIKTGSLSARSDRLAKYNQLIRIEEMLGPOANYAGRS	420

Search completed: September 29, 2005, 14:45:40
Job time : 1027 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)